On the Influence of Age of Acquisition and Proficiency on Second Language Processing

DISSERTATION

zur Erlangung des akademischen Grades

doctor philosophiae

(Dr. phil.)

im Fach Allgemeine Sprachwissenschaft

von

Juliane Domke geb. Schütte

Präsident der Humboldt-Universität zu Berlin
Prof. Dr. Jan-Hendrik Olbertz

Dekanin der Philosophischen Fakultät II
Prof. Dr. Helga Schwalm

Gutachterinnen/Gutachter: 1. Prof. Dr. Sophie Repp
2. Prof. Dr. Rainer Dietrich

eingereicht am 01. April 2015

verteidigt am 02. Juli 2015

Danksagung

Für ihre Unterstützung und hilfreichen Kommentare bei der Erstellung dieser überarbeiteten Version der Dissertation danke ich Sophie Repp und Rainer Dietrich.

Weiterhin möchte ich den folgenden Personen ganz herzlich danken:

Joanna Blaszczak, Berry Claus, Yasmin Dalati, Ingo Fehrmann, Felix Golcher, Peter beim Graben, Carola Hochfeld, Rainer Kniesche, Thomas Pinkpank, Carsten Schliewe, Katharina Spalek, Nicole Stutterheim und allen Versuchspersonen.

Ein besonderer Dank geht an Martin und Romy.

Abstract

This dissertation presents research on how the age of acquisition (AoA) and the proficiency level influence the processing of a second language. Three ERP studies (event related potentials) were conducted investigating on the processing of semantic incongruity, the processing of double nominative violation, and the processing of negative polarity item (NPI) licensing in German as a second language. Three central questions were in focus: (i) whether second language learners, like native speakers also show the dissociation between semantic and syntactic processing; (ii) whether and how AoA and / or proficiency influence second language processing, independently or interactively; and (iii) whether this influence appears gradual or discontinuous.

According to (i) results largely confirm previous research on this topic suggesting that second language processing is sensitive to the semantic-syntactic dissociation. Regarding (ii) outcomes suggest interactive influences between the AoA and proficiency on second language processing especially when syntactic and semantic-syntactic demands are enhanced as to the language processing system. This contributes to recent suggestions and augments the argument that relative to the activation and retrieval of neural resources associated with syntactic processing mechanisms proficiency may compensate for the impact of age of acquisition. With respect to (iii) the present outcomes strongly support a continuity approach indicating that changes in the activation and retrieval of processing mechanisms appear gradual.

Contents

1 Introduction

For a native speaker of German the neural mechanisms that are activated when reading or listening to an ill-formed sentence like (1a) are different from those that get activated when reading or listening to an ill-formed sentence like (1b).

(1)

a) Der Autor schreibt *den Stuhl* an seinen Freund

'The author writes *the chair* to his friend.'

b) Der Autor *schreiben* den Brief an seinen Freund

'The author *write* the letter to his friend.'

The former sentence is wrong since our knowledge of the world tells us that 'der Stuhl' ('the chair') is something to sit on rather than something to be written. The latter is consistent with our knowledge, but since there is just one author the verb form should be singular 'schreibt' ('writes') and not plural 'schreiben' ('write'). The neural activation differences highlight the neurocognitive dissociation between semantic and syntactic processing mechanisms in relation to language perception. Evidence for this dissociation between the processing of semantic and syntactic anomalies is consistently found in psycholinguistic research by using electro-physical measures such as event-related potentials (henceforth ERP). Broadly, ERP data as response to the processing of a semantic anomaly, as in (1a), provides a so-called enhanced N400 component. Generally, the N400 component is assumed to reflect processing mechanisms that infer the degree of difficulty to integrate new information into the previous (sentence) context (see Chapter 2.2.1.1). This N400 component reflects an ERP processing pattern that is different from the one evoked by the processing of a syntactic anomaly, as in (1b), which should elicit a so-called P600 component. The P600 component mirrors neural mechanisms associated with reanalysis or repair of an ill-formed syntactic structure ((see Chapter 2.2.2.1). Hence, differences between semantic and syntactic processing patterns are associated with the varying difficulties that the language processing system has to face relative to the neural processing mechanisms dissociating semantic processing demands and syntactic processing demands.

Against this background, second language research explores if the neural mechanisms that are activated as response to the processing of semantic and / or syntactic anomalies can be retrieved while processing a second language (L2). In other words, the research question is: Is it possible for L2 learners to anticipate upcoming information (be it semantic or syntactic) in their L2 in the same way as native speakers do in their L1? Accordingly, the following two main issues have repeatedly been emphasised: First, whether the activation of neural mechanisms, too, is sensitive to the dissociation between semantic and syntactic processing. In other words, do the brain responses of L2 learners also elicit N400 and P600 processing patterns, respectively? Second, whether those ERP processing patterns are—and if yes, to what degree—similar or different from native language processing (see below for references).

Two factors that are claimed to influence L2 processing and hence to cause differences compared with L1 processing are *age of acquisition* (i.e., the age at which someone started to learn his / her L2) and *proficiency* (i.e., the current level of lexical and grammatical knowledge in case of L2). The influence of age of acquisition (henceforth AoA) is based on the assumption that learning a (second) language is subject to maturational constrains—i.e., there is a biologically determined time-window within which (a second) language has to be acquired in order to reach native-like competence (Penfield & Roberts, 1959; Lenneberg, 1967). If L2 acquisition starts within this time-window, one may reach native-like competence. This is usually referred to as early L2 acquisition. Late L2 acquisition, then, starts beyond this critical time-window and relates to the assumption that native-like competence is no longer achievable. There is much debate about whether there exists such a biologically determined time-window, and what exact age determines the border between early and late L2 acquisition. For insightful discussions, the reader is referred to Abrahamsson and Hyltenstam, 2009; Birdsong, 1999, 2006; DeKeyser, 2000; Eubank and Gregg, 1999; Long, 1990, 2005; Hyltenstam and Abrahamsson, 2003; Singleton, 2005; and, more recently, Meisel, 2011. See also

Mũnoz and Singleton, 2011; Vanhove, 2013; and Birdsong, 2014 for recent reviews on this issue. The present thesis will not focus on the question whether there exist maturationally constrained periods that restrict L2 acquisition. It is exclusively interested on how (—and not if—) AoA influences the processing of an L2.

The proficiency of an L2 learner indicates a specific level of knowledge acquired in the L2 and is commonly referred to as his / her lexical and grammatical L2 competence. There is the common idea inferring that higher proficiency reflects higher language competence. There are several ways to measure the proficiency of an L2 learner and, hence, to classify the beginning, intermediate, and advanced stages of L2 learning and knowledge (for overviews, see Leclercq & Edmonds, 2014; Thomas, 1994). Nonetheless, proficiency in L2 research is a rather vague concept that lacks a clear and consistent definition (cf. Hilton, 2014, pp. 27). A comparable vagueness appears with the determination of lexical and grammatical L2 competence. I.e., the European Union has established the *Common European Framework of Reference for Languages* (CEFR), which includes the classification of (L2) language competence into six stages: A1 = breakthrough, A2 = waystage, B1 = threshold, B2 = vantage, C1 = effective operational proficiency, and C2 = mastery (cf. Council of Europe, 2011, p. 23). Note that the descriptions of the features of (L2) language competence given by the CEFR rely on communicative rather than on linguistic skills (ibid, Chapters. 4–5). Still, these categories may also be applied to mere linguistic competence that can be measured by standardized proficiency tests such as the C-Test (Klein-Braley & Raatz, 1982; see also Grotjahn, Klein-Braley, & Raatz, 2002). The C-Test is a cloze test which is based on the principle of reduced redundancy. This principle assumes that the less (lexical and grammatical) information is needed in order to accurately fill in the gaps, the higher the lexical and grammatical knowledge of a language. For a recent extensive bibliography on the C-Test, see Grotjahn (2014). An example of a C-Test is given in appendix 9. The present thesis refers proficiency to the level of lexical and morpho-syntactic L2 competence as it can be measured by the standardized C-Test. Communicative L2 competence in terms of the CEFR will not be issued. Moreover, in the present thesis, lexical L2 competence will be associated with the L2 learner's vocabulary knowledge, while grammatical L2 competence will be narrowed down to the morpho-syntactic knowledge of an L2 (see also above).

One major challenge faced by psycholinguistic investigations within the field is that there is an obvious correlation between the two factors AoA and proficiency, i.e. later AoA indicates lower proficiency (for overviews see, e.g., Birdsong, 2014; Mũnoz & Singleton, 2011; see also Bialystok & Hakuta, 1999). Traditionally, studies have reported on the basis of behavioural evidence that the AoA of an L2 learner reliably predicts his / her proficiency level (Birdsong & Molis, 2001; Flege, Yeni-Koshmian, & Liu, 1999; Johnson & Newport, 1989). For instance, Flege et al. (1999) report the accuracy of grammaticality judgements of correct and incorrect English sentences for English native speakers and L2 learners of English (with L1 = Korean). L2 learners' AoA ranged between one to 23 years.[1] As is illustrated in Figure 1.1 below, their results reveal a strong correlation between AoA and the accuracy scores elicited by L2 learners. Interestingly, next to the general decrease of mean accuracy scores, there is also an increase of variance for these scores as a function of the growing AoA. As can be seen in Figure 1.1, an L2 learner with AoA = 12 may score just above chance level, while another L2 learner with AoA = 12 scores above 90%. In this specific case, AoA cannot account for the differences between mean accuracy scores and rather the level of proficiency may determine the difference between the outcomes of these two L2 learners. Hence, proficiency may be treated as an independent factor rather than the dependent variable. Likewise, the level of L2 proficiency determines the accuracy of grammaticality judgements. Still, the treatment of AoA and proficiency as independent influencing factors does not neglect their inherent correlation.

[1] Flege et al. (1999) use the term 'Age of Arrival' (in the USA) rather than 'Age of Acquisition'. In this thesis, AoA is explicitly referred to as the term 'Age of Acquisition'—i.e., the age at which an L2 learner starts to get exposed to L2 input on a frequent basis.

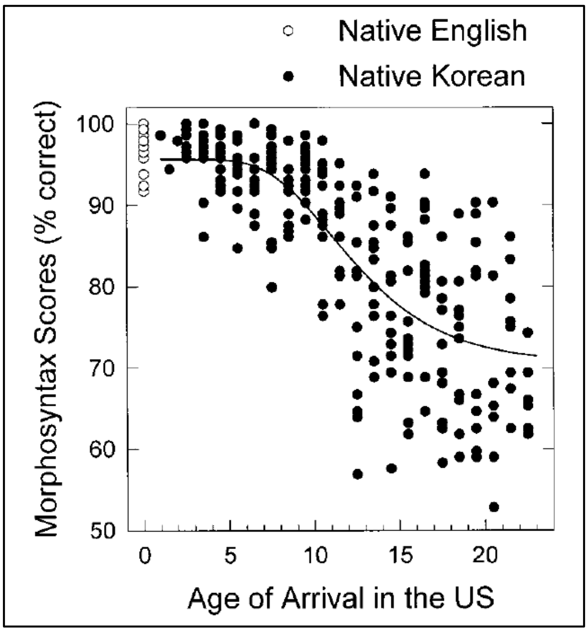


Figure 1.1: Adapted from Flege et al., (1999, p. 85): Mean accuracy in the grammatical judgment task performed by native speakers (transparent dots) and L2 learners (black dots) of English. Y-axis displays mean accuracy scores for each participant; AoA in years is mapped on x-axis. Original citation: "Fig. 2: The grammaticality judgment test scores obtained for 24 native English and 240 native Korean participants. The data for the 240 Koreans have been fit to the Gompertz-Makehm distribution (solid line)."

Concerning ERP data as response to L1 and L2 processing, previous research reports respective differences between the according brain responses. Furthermore, the differences appear more pronounced for syntactic than semantic processing (see Chapter 2.2; for comprehensive reviews on ERP differences between L1 and L2 processing, see also Hahne, 2000; van Hell & Tocowitz, 2010; Moreno, Rodriguez-Fornells, & Laine, 2008; Mueller, Rueschemeyer, & Friederici, 2006; and very recently Steinhauer, 2014). Despite the nature of processing, including the degree of variation, underlying reasons for those differences have been attributed to the impact of both L2 learners' AoA and proficiency levels. Furthermore, attempts have been made to find singular causality given the two influencing factors so that studies have generally accounted for their inherent correlation mainly by controlling for either one (AoA) or the other factor (proficiency). That is, in order to check whether AoA has impacts on L2 processing, it is necessary that L2 proficiency remains constant. Or the other way around, when investigating possible proficiency influences, L2 learners need to have similar AoA. With respect to experimental designs and statistical data analyses, this has been realized by separating L2 learners in accordance with their AoA (e.g., early vs. late AoA) or proficiency (e.g., low vs. high) and hence categorizing factors (e.g., by relying on the mean AoA or mean proficiency) which are actually continuous (for references and discussions see Chapter 2.2). Only recently, suggestions have been made to treat the factor of investigation (e.g., proficiency) continuous rather than categorical to still control for the other factor (e.g., AoA) (Newman, Tremblay, Nicols, Neville, & Ullman, 2012; Tanner, Inoue, & Osterhout, 2014; van Hell & Tanner, 2012; similarly see discussions in Chapters 2.2.1.3 and 2.2.2.4; see also Baayen, 2010 for general criticism of data analysis based on simple factorial / categorical statistics). In addition to the continuous treatment of the influencing factors, there is also recent evidence for an interactive influence of the factors, AoA and proficiency, on L2 grammatical development. Native-like processing is approached in a step-wise manner as a function of improving L2 proficiency when AoA is held constant, which suggests continuity (McLaughlin et al, 2010; Osterhout, McLaughlin, Pitkanen, Frenck-Mestre, & Molinaro, 2006; Steinhauer, White, & Drury, 2009; Tanner, McLaughlin, Herschenssohn, & Osterhout, 2013; see also Osterhout, McLaughlin, Kim, Greenwald, & Inoue (2004)).

This thesis takes up these latter suggestions and presents research on how AoA and proficiency have an influence on semantic and syntactic L2 processing by treating both as continuous factors. Moreover, these two factors are not observed independently, but their potential interaction is integrated into the analysis. This is statistically sensitive since both these factors are correlated, and the treatment of correlated factors as independent variables and their multiplication to an interaction limits the statistical interpretation, which, in turn, results in restricted interpretations on the basis of model criticism (for more detailed considerations see Chapter 3.4.1). However, the very nature of this correlation and its influence on L2 processing are of particular interest. More precisely, the present thesis addresses three ERP studies that investigate the potential differences between the ERP patterns elicited by native speakers and L2 learners of German. It further examines the impact of AoA and proficiency on potential processing differences and additionally scrutinizes whether these influences appear to be independent or interactive. Since both factors are included in the analysis, it will be interesting to see whether the (statistical) weightings of both factors vary and whether one over the other better explains the present data. It will be referred not only to the distinction between L1 and L2 processing (e.g., which factor more accurately predicts potential differences between L1 and L2 processing) but also to the dissociation between semantic and syntactic processing (e.g., do the factors make same or different predictions with respect to the linguistic structure that is processed). The materials used in the three ERP studies vary in accordance with the linguistic structure and consequently with the demands regarding the language processing system (i.e., the neurocognitive sources). Those are semantic processing demands triggered by a semantic anomaly, syntactic processing demands prompted by double nominative violation, and a combination of syntactic-semantic processing demands generated by the licensing failure of a negative polarity item (henceforth NPI[2], see also Chapter 2.2.3). In general, the results go in line with previous findings suggesting the following two things: On the one hand, L2 learners, like native speakers, reveal the neural dissociation between semantic and syntactic processing mechanisms. On the other hand, the native and non-native processing patterns show differences. Furthermore, the degree of such differences is primarily determined by the anomaly of the linguistic structure that is processed—i.e., differences between L1 and L2 ERP responses to semantic processing are less striking than to syntactic processing. Beyond this, the present results indicate that both AoA and proficiency influence the L2 processing patterns, and that AoA is generally more predictive in explaining the data. However, in syntactic L2 processing proficiency influence gains importance with respect to the ability to retrieve the associated neural mechanisms, while in combined semantic-syntactic L2 processing retrieval of semantic mechanisms are affected by this very character of the interaction of the two factors. Finally, changes in the strength of impact (including the statistical weighting) of both influencing factors appear to be gradual

[2] NPIs like German 'jemals' ('ever') are linguistic elements whose occurrence in sentences is restricted to semantic and syntactic licensing conditions. There has to be an appropriate licensor, and it has to be structurally accessible to the NPI (for a more detailed description and references, see Chapter 2.2.3).

rather than discontinuous, which supports the recent idea to treat both variables continuously and not categorically.

The present thesis is structured as follows:

Chapter 2 introduces the ERP method (2.1). It further reviews the important ERP findings elicited by L1 and L2 comprehension relative to the structures chosen for the present investigations on potential differences between L1 and L2 processing, and with respect to the varying neural mechanisms due to AoA and / or proficiency influence. Those are semantic incongruity (2.2.1), syntactic double nominative violation (2.2.2), and semantic-syntactic licencing violation of an NPI (2.2.3). Chapter 2.3 takes up the formerly reviewed findings and introduces a neurocognitive model that accounts for varying differences between L1 and L2 processing. Chapter 3 covers the detailed description of experimental routines, including participants, stimulus material presentation, experimental procedures, and also methods of statistical analyses. Chapter 4 comprises the presentation and interpretation of the results of the three datasets. Finally, Chapter 5 presents a general discussion which includes conclusions that are drawn by evaluating the results of the present investigations in the light of the design and analysis employed and with respect to former results reviewed in Chapter 2.2.

2 ERPs as Response to L1 and L2 Processing

Chapter 2 introduces the experimental method used for the present investigations on German L1 and L2 processing, namely ERP. The design of the present study is based on the violation paradigm. For each of the three German structures in focus, viz. semantic (in-) congruity (see examples in [2] below), syntactic double nominative violation (see examples in [3] below), and semantic-syntactic NPI-licensing structures (see examples in [4] below), the stimulus material consists of sentences entailing two conditions that are directly compared with each other. To be more precise, the subjects of comparison are the so-called critical items of (i) a control condition (correct / congruent, see also sentences [a] in examples [2]–[4]) and (ii) a deviant condition (incorrect / incongruent, see also sentences [b] in examples [2]–[4]).[3]

(2)[4]

a) Der$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ schreibt *den$_{ACC}$ Roman$_{ACC}$* und gewinnt einen$_{\mathbf{ACC}}$ Preis$_{\mathbf{ACC}}$.
'The$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ writes *the$_{ACC}$ novel$_{ACC}$* and wins a$_{\mathbf{ACC}}$ prize$_{\mathbf{ACC}}$.'

b) *Der$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ schreibt *den$_{ACC}$ Stuhl$_{ACC}$* und gewinnt einen$_{\mathbf{ACC}}$ Preis$_{\mathbf{ACC}}$.
*'The$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ writes *the$_{ACC}$ chair$_{ACC}$* and wins a$_{\mathbf{ACC}}$ prize$_{\mathbf{ACC}}$.'

(3)

a) Der$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ schreibt *den$_{ACC}$ Roman$_{ACC}$* und gewinnt einen$_{\mathbf{ACC}}$ Preis$_{\mathbf{ACC}}$.
'The$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ writes *the$_{ACC}$ novel$_{ACC}$* and wins a$_{\mathbf{ACC}}$ prize$_{\mathbf{ACC}}$.'

b) *Der$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ schreibt *der$_{NOM}$ Roman$_{NOM}$* und gewinnt einen$_{\mathbf{ACC}}$ Preis$_{\mathbf{ACC}}$.
*'The$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ writes *the$_{NOM}$ novel$_{NOM}$* and wins a$_{\mathbf{ACC}}$ prize$_{\mathbf{ACC}}$.

---

[3] Throughout the thesis critical items given in examples are in italics; '*'-indexed sentences are incongruent / incorrect

[4] NOM = Nominative; ACC = Accusative.

(4)

a) Kein$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ hat den$_{\mathbf{ACC}}$ Roman$_{\mathbf{ACC}}$ *jemals* geschrieben.
'No$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ has the$_{\mathbf{ACC}}$ novel$_{\mathbf{ACC}}$ *ever* written.'

b) *Der$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ hat den$_{\mathbf{ACC}}$ Roman$_{\mathbf{ACC}}$ *jemals* geschrieben.
*'The$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ has the$_{\mathbf{ACC}}$ novel$_{\mathbf{ACC}}$ *ever* written.'

The structures were chosen in accordance with their processing demands. The processing of semantic incongruity within a German sentence like 'den Stuhl' ('the chair') in (2b) is supposed to enhance the activation of neural mechanisms underlying lexical-semantic resources (see Chapter 2.2.1). Sentences with double nominative case-markings, as in (3b), indicate a structural violation because the nominative case-marking on the second NP is ungrammatical in German. Such misconstructions should present processing difficulties that are reflected by the devotion of neural mechanisms associated with syntactic processing (see also Chapter 2.2.2). The processing of a structure entailing inappropriate NPI licensing, as in (4b), has been chosen since both semantic and syntactic processing demands are enhanced (see Chapter 2.2.3).

For the three structures in focus, the processing patterns based on ERP data elicited by native German speakers are well documented (see Chapters. 2.2.1.1, 2.2.2.1 and 2.2.3.1 below). ERP data elicited by L2 learners, which reflect the processing of semantic incongruity and double case violations, has been reported quite sufficiently. Yet, with respect to ERP differences between native speakers and L2 learners, the results and especially their interpretations in the light of AoA and proficiency influence are rather diverse, as will be seen below (Chapters 2.2.1.2 and 2.2.2.3). ERP data reflecting the L2 processing of NPIs is scarce (see Chapter 2.2.3.2 below). Before reviewing the L1 and L2 processing data of these three structures, a short introduction to the ERP method is given.

2.1 ERP Method

Over the past 30 years the ERP method has become a major tool for the study of language processing. This method offers a high temporal resolution of the cortical processing (re)action, contributing to the fact that language processing can be observed as it happens (for a coherent introduction to the ERP method / technique, see Luck, 2005). The ERP method provides a detailed analysis of the so-called ERP components, thereby allowing a close eye on exact points in time (or time-periods) when and how[5] specific language phenomena are processed (for general overviews of ERP components relative to language processing, see Drenhaus & beim Graben, 2012; Garnsey, 1993; Kaan, 2007; and more recently, Morgan-Short & Tanner, 2014). The ERP method is based on EEG conduction, which is a non-invasive, direct measurement of excitatory postsynaptic brain potentials that are located on the apical dendrites. The EEG is associated with electrical brain activity of larger cortical neuron-assemblies and what is measured—i.e., what makes up the signal—is the repeatedly simultaneous discharge of these cortical neuron-assemblies (cf. Christian, 1984, pp. 3–4; for a detailed description of the sources of underlying brain potentials measured by EEG, see also Zschocke & Hansen, 2011, Chapters 1 and 2). Substantially, these signals are frequency outcomes. Potential changes in frequency are dependent on cortical activity, which means that the signals may have different characteristics.

The cortical activity is conducted by electrodes that are placed on the scalp via an electrode-cap. The pattern in which the electrodes are located on the scalp is implemented corresponding to the appropriate and widely accepted standardized distribution prescribed by the American Electroencephalographic Society (Sharbrough, et al., 1995; see also Figure 2.1 below).

[5] The way of how language is processed in a technical way with regard to the analysis of components tells about the shape and quality of the component at hand (cf. Garnsey, 1993, pp. 344–345).

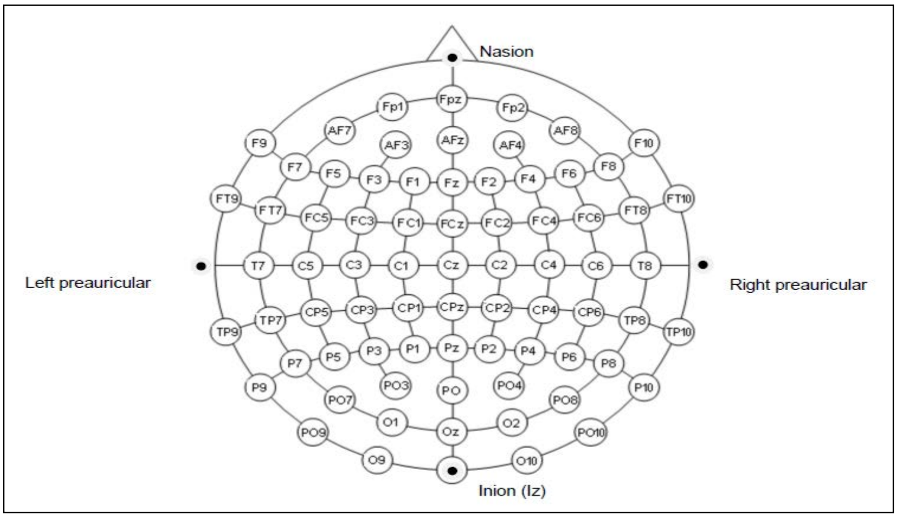


Figure 2.1: Electrode positions according to the standardised 10-20 system of the American Electroencephalographic Society. Retrieved from csus-dspace.calstate.edu.

The EEG outcome—i.e., the measured stream of changes in cortical activity—is large by means of the dataset to be analysed. Yet, only specific parts of the data stream are relevant to further investigation. These relevant parts directly link with the measured electrical activity as response to the processing of a specific experimental manipulation (e.g., the critical item(s) within the stimulus material). Critical items within the stimulus material have to be presented repeatedly by exact points in time, the so-called *events*. The matching events, then, are averaged to an ERP—the repeated and time-locked EEG parts of the task / material at hand are taken off the entire raw EEG and subsequently are averaged for (each) participant, and, finally, for the whole group of participants (i.e., grand average ERP). The averaging technique is necessary for various reasons.[6] Most of all, EEG signals are noisy, indicating that not only cortical activity related to experimental stimuli but also irrelevant activity is conducted. Especially facial and / or body movements as well as outside electrical noise (power plugs, amplifier, light source, etc.) influence the signal-to-noise-ratio. This activity is undesired for experimental

[6] Next to the averaging technique, there are alternative ways for analysing ERP data; Dambacher, Kliegl, Hofmann, and Jacobs (2006) analysed single EEG trials. Drenhaus, beim Graben, Saddy, and Frisch (2006) introduced ERP data elicited by the Symbolic Resonance Analysis (beim Graben & Kurths, 2003), where EEG background noise of the critical ERP period may be used to dissociate those ERP components that are not elicited by the averaging method.

purpose and therefore considered an artefact. Artefacts have to be removed or rather averaged out of the data. Next to the averaging process, further instructional and filtering options may be applied to erase artefacts.[7]

Typically, a grand average ERP wave is visualized as a voltage-time curve. The voltage-time curve commonly depicts a 1,000 ms (+ / - 500 ms) time-window post stimulus and can be analysed for its ERP components. According to Garnsey (1993), an ERP component '[…] is part of the waveform that is consistently sensitive to particular kinds of experimental manipulation, and thus thought to be the manifestation of a particular process or set of processes' (pp. 344–345). Commonly, two or more ERP components are contrasted, and this contrast is inferred from a violation paradigm (see above). Such contrast comprises the ERP waveform of a control condition, which is compared with an analogues ERP waveform of the deviant condition. A voltage-time curve can be seen in Figure 2.2 on left-hand side. The solid black line depicts the ERP response to the critical item of the control condition (the word 'sweet'). The dashed blue line illustrates the ERP signature in accordance with the critical item of the deviant condition (the word 'anxious'). The difference between the ERP components of both conditions is described and characterized as the ERP effect. The contrasted ERP waves and their differences (ERP effects) are analysed in terms of specific parameters, most commonly polarity, time, and distribution. Polarity is a binary parameter; the ERP signal may either be positive or negative, depending on the baseline x-axis, which corresponds to 0 µV—i.e., ranging from approximately -5 µV to +5 µV.[8] The time-parameter links to the exact point in time when the signal has its maximum amplitude (peak), or starts to move away from the baseline (x-axis), or, in case of the deviant ERP wave, away from the wave of the control condition (onset). For example, see again the left-hand side ERP in Figure 2.2: Between 300 and 500 ms, both conditions show a negative-going amplitude reaching its maximum (i.e. peak) around 400 ms; this ERP component is known as

[7] There still might remain some sort of brain activity that could be considered artefacts (i.e., resulting from eye-movements such as blinks), which cannot simply be cancelled out by averaging. These artefacts have to be removed manually or automatically by the technical advice in accordance with the programs used for processing raw data.
[8] Conventionally, in an ERP graph negative voltages are plotted up and positive ones are plotted down.

N400 (N = negativity; 400 = peak; for a detailed discussion on the N400 see Chapter 2.2.1.1 below). Thus, polarity and time values are precisely observable and contribute to the determination of an ERP component. They also cover the name-giving function. Furthermore, the amplitude of the N400 in the deviant condition is more enhanced than in the control condition. This very difference between both N400 amplitudes determines the ERP effect (here: N400 effect, again, see Chapter 2.2.1.1 for detailed descriptions).

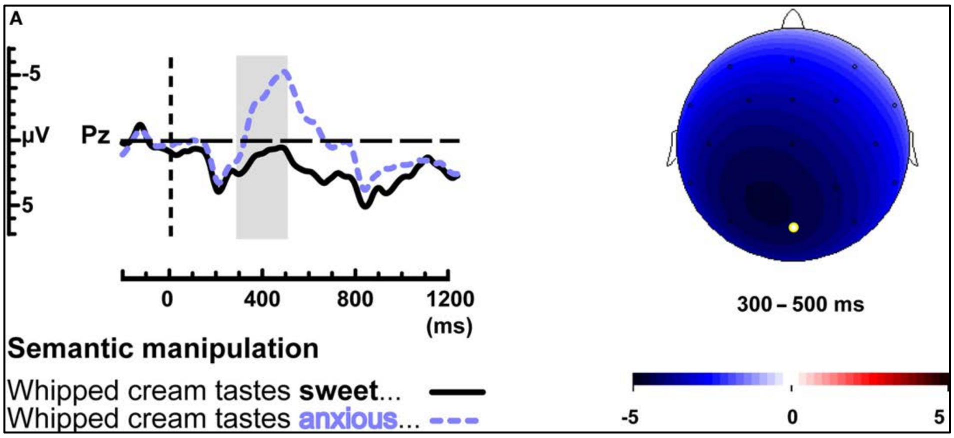


Figure 2.2: Adapted from Kos, Vosse, van den Brink and Hagoort (2010, p. 5): Voltage time map (left) and voltage difference map (right) of an N400 effect elicited by semantic manipulation. Original citation: "Figure 1. (A) Grand-average waveforms for ERPs elicited by the semantic anomalies (dotted, blue line) and their correct controls (solid, black line) for electrode Pz and the scalp distribution of the N400 effect elicited by the semantic manipulation between 300 and 500 ms after critical word onset. ([…] [T]he waveforms are time-locked to the onset of the critical word (0 ms) and negative voltage is plotted upward. […])".

An ERP effect may also be illustrated as a voltage difference map. On the right-hand side, in Figure 2.2 (see above), the voltage difference map illustrates the average negativity[9] (in μV) between 300 and 500 ms. The negativity appears stronger on the posterior than anterior electrodes, indicating mean voltage differences as to scalp distribution or topography. The parameter scalp distribution therefore focuses on the local strength of the ERP effect—i.e., whether the observed differences between two ERP components are equally distributed over the entire scalp or whether there are

[9] The voltage difference is calculated as follows: the voltage average (μV) of 'sweet' is subtracted from the voltage average (μV) of 'anxious'.

differences in the strength between scalp-sites (anterior / posterior) and / or scalp-hemispheres (right, central or left).

With respect to the statistical analyses of ERPs, usually a specific time-window is derived in accordance with the strength of the observed ERP effect on the basis of visual inspection. Relative to this specific time-window, the mean voltages of each electrode averaged for each participant and for each condition are determined and subject to statistical analysis. In recent years, the methods to statistically analyse ERP data have progressed mainly due to linear mixed-effects modelling techniques (e.g., Baayen, Davidson, & Bates, 2008). The performance of linear mixed-effects models on ERP data has some general advantages: it relatively easy handles missing data and (non-) sphericity assumptions.[10][11] Chapter 3.4.4 offers a detailed description of the models performed to analyse the upcoming ERP datasets, and thereby explains the above mentioned advantages more precisely.

In summary, the ERP method has been a useful tool to investigate language processing phenomena online. ERP data offers a very high temporal resolution and can be characterized for specific ERP components. Usually, it is the difference(s) between two (or more) ERP components that are of interest, as this very difference determines an ERP effect. In what follows, two ERP effects associated with language processing on the basis of single sentence contexts, viz. N400 effect and P600 effect, are reviewed and described in more detail.

[10] Handy, Nagamatsu, Mickleborough, & Liu-Ambrose (2009) offer a thorough discussion on the assumption of sphericity and its concerns in connection with ERP data.

[11] Still, the advantages of performing mixed-effects models on averaged ERP data are limited (compared with behavioural data) since single trial variation is not available for being included as a random effect (J. Verissimo, personal communication, November 03, 2014, see also Chapter 3.4.3).

2.2 Language-Specific ERP Components and ERP Effects

As already indicated in the introduction, the field of (single) sentence comprehension traditionally assumes two broadly distinct processing mechanisms, namely lexical-semantic processing mechanisms and syntactic processing mechanisms.[12] On the basis of experimental manipulations the dissociation of processing mechanisms is reflected by two well-established ERP components and their resulting effects, namely N400 and P600, respectively. Both ERP components are reviewed in the next sections with respect to L1 and L2 processing, and in the light of the materials and linguistic violation paradigms used in the upcoming experiments of the present thesis.

2.2.1 Semantic Incongruity

2.2.1.1 N400 and L1 Processing of Semantic Incongruity

In this thesis, semantic incongruity processing is issued at the single sentence level. Incongruity is indicated by a content word at a sentence's medial position, which does not match the context of the prior sentence, given as 'der Stuhl' ('the chair') in (2b), as seen above. A comparison between the ERP of incongruent 'the chair' (2b) and the ERP of congruent 'the novel' (2a) should reveal the so-called N400 effect. As briefly indicated in Chapter 2.1, the N400 effect infers the comparison of two (or more) N400 components—i.e., enhanced negativities with a maximum peak around 400 ms post stimulus. The difference is usually the largest over posterior scalp-sites. The term 'N400 component' is not synonymous to the term 'N400 effect'. While the 'N400 component' is evoked by the processing of any (content) word, the term 'N400 effect' is associated

[12] Alongside, there are further linguistic research fields that use ERP, such as phonology (see e.g., Molfese, Key, Maguire, Dove, & Molfese, 2008 for an overview of ERPs in speech perception) or pragmatics (see van Berkum, van den Brink, Tesink, Kos, & Hagoort, 2008 for data on pragmatic processing). Those will not be discussed any further since the present investigations solely focus on semantic and syntactic processing.

with the mean voltage difference between the ERP of the incongruent condition (context mismatch) and the ERP of the congruent condition (context match) within the relevant time-window (for comprehensive reviews on the N400 effect, see Kutas & Federmeier, 2000, 2011).[13] The N400 effect was first reported by the seminal studies of Kutas and Hillyard (1980a, 1980b). They found that the N400's amplitude is enhanced when a semantically incongruent / unexpected word indicating a semantic violation—relative to a semantically congruent / expected word—is presented at the end of a sentence. The incongruent word cannot easily be integrated into the former sentence context—e.g., 'dog' in 'I take my coffee with cream and *sugar* / **dog*'. Furthermore, it was shown that given the context of the former sentence, the expectation of a word to come up next—i.e., the word's cloze probability—correlates with the size of the N400 amplitude (Kutas & Hillyard, 1984). Hence, cloze probability puts forward a certain prediction for a forthcoming word while processing a sentence. Accordingly, the larger is the N400 component, the lesser prediction is met (Federmeier, 2007; again see Kutas & Hillyard, 1984; see also van Petten & Luka, 2012). Moreover, this cloze probability account does not exclusively rely on the context of the former sentence, which indicates a semantic violation, but also depends on world knowledge. Likewise, the amplitude of the N400 component is enhanced when an upcoming word yields a sentence interpretation which does not match world knowledge—e.g., 'Dutch trains are *yellow* / **white*' (see Hagoort, Hald, Bastiaansen, & Petterson, 2004; see also Hagoort, Baggio, & Willems, 2009).

Explanations concerning the functional role of the N400—i.e., which underlying cognitive process(es) or state(s) does it represent—are still a matter of debate.[14] According to language processing mechanisms, the most common view relates the N400 to the activation of mechanisms reflecting lexical-semantic integration processes.

---

[13] The remainder of this thesis refers to the term N400 as the N400 component. The N400 effect will be associated with the significant difference between two (or more) N400 components within a violation paradigm.

[14] Van Petten and Luka (2006) offer a thorough discussion on the neural processes that are assumed to be related to the N400. Lau, Phillips, and Poeppel (2008) suggest a neuroanatomical model by localizing those neural processes that may reflect the N400 components.

Given the assumption that sentence processing is highly predictive, and that specific predictions are made up by the (former) sentence context and / or our world knowledge (see above), the integrative account suggests that the N400 represents a reliable ERP component concerning the cortical reflection of lexical and semantic integration mechanisms (e.g., Friederici, 1995, 2002; again see reviews by Kutas & Federmeier, 2000, 2011). It means that the processing of a sentence yields updating mechanisms with the new incoming information (i.e., a word). Difficulties with the updating processes emerge if the new incoming information is less or non-predicted on the basis of the former sentence context and / or world knowledge. Consequently, the more enhanced appears to be the amplitude of the N400, the more difficult seem the updating processes (but see e.g., Brouwer, Fitz, & Hoeks, 2012; Deacon et al., 2004; and also Kuperberg, 2007 for alternative explanations of the functional role of the N400). Given the N400 as a lexical-semantic ERP component, it is assumed to occur independent of—though parallel to—syntactic processing components (again see e.g., Friederici, 1995, 2002; see also Steinhauer & Conolly, 2008; also see Brouwer, et al., 2012; or Kuperberg, 2007 for different accounts). Yet, there is evidence that within a semantically incongruent context additional syntactic violation of the structure may further enhance the N400 (Hagoort, 2003; Osterhout & Nicol, 1999; Ye, Zhan, & Zhou, 2007).

2.2.1.2 N400 and L2 Processing of Semantic Incongruity

Taking up the integrative account of the N400 previous outcomes suggests that the corresponding neural mechanisms are robustly operative in L2 processing. In other words, L2 learners in their L2 anticipate the upcoming semantic information and retrieve neural mechanisms associated with updating processes. With regard to ERP data, it means that L2 processing patterns evoked by semantic incongruity reliably reveal N400 effects (see below). Yet, differences between N400 effects elicited by L2 learners' ERPs compared with those of native speakers have been reported with regard to strength, latency, and distribution. However, these differences have somewhat inconsistently been claimed either to be due to varying AoA or proficiency of the L2 learners. Therefore, the occurrence of potential differences between L1 and L2 N400 effects is not clearly predictable. In what follows, a selective overview is given on ERP data elicited by L2 learners when processing semantic incongruities.

In an ERP study, Ardal, Donald, Meuter, Muldrew, and Luce (1990) investigate the processing of semantically congruent and incongruent English and French sentences like 'I generally like menthol **bottles*' (visually presented), as obtained for native English speakers and L2 learners of English (L1 = French). The L2 learners are grouped on the basis of their AoA in their L2 (early <11, late >11)[15] and all are highly fluent (i.e., highly proficient) in their L2. Their ERP results do not show any significant differences in the N400 amplitude or latency between L2 learners with early AoA or late AoA. The authors further compare the ERPs of monolinguals (L1) and bilinguals in their L1 and L2. They do not control for individual L2 learner's AoA and take into consideration that the initial analysis did not reveal any differences. ERPs of all three groups—monolinguals and L2 learners in their L1 and L2—show a strong negativity for incongruent words compared with congruent words (i.e., N400 effect). The differences in the N400 effects between the three groups are determined by a longer latency (approximately 40 ms), frontally reduced amplitudes, and more left-lateralized peaks observed for L2 learners in their L2, compared with their L1 and also compared with monolinguals. Ardal and his colleagues interpret these differences as indication of less automatized processing, assuming that semantic operations underlying the N400 reflection are processed more slowly in the L2, which in turn indicate higher processing costs. They account this quantitative difference between L1 and L2 processing for L2 fluency (i.e., L2 proficiency) influence and put forward that, in general, L1 fluency is higher than L2 fluency, which suggests a more automatized activation and retrieval of updating mechanisms.

A delay in the N400 amplitude, as response to semantic incongruent words elicited by L2 processing, is also reported by Weber-Fox and Neville (1996). They attribute their

[15] The distinction between early and late AoA is directly adopted from Ardal et al. (1990).

finding to AoA influence. They investigate semantic processing of L2 learners of English (L1 = Chinese) with different AoA. The authors rank the L2 learners into five groups, accordingly (AoA=1–3; 4–6; 7–10; 11–13; >16).[16] Weber-Fox & Neville additionally obtain the number of years to English (L2) exposure for the L2 learners, their self-rated proficiency in reading, writing, listening and speaking in both languages, and L2 learners' general amount of L1 / L2 (Chinese / English) usage at home and at work / university. The authors further conduct a standardized English proficiency test. Between-group comparisons reveal that the number of years of L2 exposure decreases as AoA increases, except between the groups 11–13 and >16 who report similar amounts of L2 exposure. The results of the L2 learners' self-ratings show that the groups 1–3, 4–6, and 7–10 sense themselves to be more proficient in their L2 than in their L1. L2 learners in the 11–13 group rate themselves equally proficient in both languages (except for writing). The >16 group feels far more proficient in their L1 considering all four language skills. Results of the standardized proficiency tests reveal significantly lower scores for the groups 11–13 and >16 relative to monolingual results (see Footnote 16). Weber-Fox & Neville do not include the proficiency scores as a variable in any further analysis. They heavily rely on AoA and years of exposure as the possible predictors for between-group differences. Behavioural and brain responses of semantic processing are conducted by semantically congruent and incongruent sentences such as 'The scientist criticised Max' *proof* / **event* of the theorem'. Behavioural results show lower accuracy rates only for the >16 group especially for the semantically incongruent condition. The authors do not report rating differences for the congruent condition in comparison with the remaining L2 learner groups. ERPs of all of the L2 learner groups show an N400 effect. However, Weber-Fox & Neville report latency and distributional differences of the N400 effects between groups compared with the results of their former study with English monolinguals (again, see Footnote 16). Mean brain responses (i.e., N400 effect) elicited by the 1–3 group do not differ

[16] The authors do not conduct monolingual data simultaneously. For comparison with native speakers' ERPs, they call on the results of an earlier study of Neville, Nichols, Barrs, Forster, and Garrett (1991), where the same stimulus material was presented to native English speakers, and their behavioural and brain responses were recorded.

from those elicited by the monolinguals. Both reveal an N400 effect showing the typical posterior distribution with a right-lateral bias. The enhanced N400s relative to incongruent sentences obtained for the groups 4–6, 7–10, and 11–13 also reveal a posterior distribution but no hemispheric differences. The N400 is equally strongly distributed over both cortical hemispheres. Likewise, the N400 effect elicited by the >16 group does not show any hemispheric differences. Yet, the >16 group shows a significant N400 effect within a broader time-window, which additionally appears to be less strong on the posterior electrodes compared with the other L2 groups. Weber-Fox & Neville further report significant delays in peak latency of the N400 amplitude as response to the incongruent condition for the groups 11–13 and >16 relative to monolinguals, but not when compared with the remaining groups with earlier AoA. The mean peak latency occurs with a delay of 23 ms for both of the former groups (again compared with the monolinguals' mean peak latency). Further correlation tests show that the latency shift is significantly predicted by both increasing AoA and decreasing years of exposure. Weber-Fox & Neville interpret their results to the effect that the N400 as reflection to semantic incongruity processing is sensitive to AoA in that the latency is prolonged when AoA starts at age 11 or later. They further conclude that generally, with increasing AoA, the N400 effect appears to be more left-lateralized (loss of hemispheric differences) and reduced in strength. Nevertheless, Weber-Fox & Neville acknowledge the fact that proficiency scores differed in the majority of tests once AoA was >10, when compared with participants with AoA <10. They point out that an interpretation towards clear AoA evidence may be limited. It also might be possible that the between-group differences in the N400 amplitudes are due to L2 proficiency differences.

Hahne (2001) investigates the auditory processing of semantic incongruity by native German speakers and L2 learners of German (L1 = Russian) with AoA >10. Semantic congruity and incongruity is presented on the final participle of German sentences such as 'Die Tür wurde *geschlossen*' ('The door was being *closed*') and 'Der Ozean wurde **geschlossen*' ('The ocean was being **closed*'), respectively. Behavioural results reveal less accurate judgements and longer reaction times for L2 learners, when compared with native speakers. ERP results show an N400 effect for both native speakers and L2 learners. The N400 effect elicited by the L2 learner group appears to be attenuated.[17] Interestingly, Hahne indicates that this attenuation is due to the enhanced N400 of the congruent condition (when compared with the L1 group). For her, it seems that the integration of words into the prior context is more demanding at late AoA *per se* and that the integration problems of semantic violations are not affected in terms of strength. Although she does not report any *post hoc* analysis, Hahne acknowledges the high correlation between AoA and proficiency. She measured the proficiency by self-evaluation of the L2 participants. The outcomes show a rather high mean score, even though single scores between subjects largely vary.[18] Further settings such as mean years of instructed exposure to the L2 (two–168 months) and mean residence in Germany (12–204 months) are also reported. They, too, show a rather wide range and thus may confound with AoA. Therefore, it is not clear which factor predicts the reported integration difficulties revealed by differences of the strength of the N400 effect between L1 and L2 processing patterns, even though they are relatively small (for related comments, see Moreno & Kutas, 2005, p. 207).

Ojima, Nakata, and Kagigi (2005) account for behavioural and ERP results of high and low proficient L2 learners of English (L1 = Japanese) compared with the brain responses of native English speakers. L2 learners' AoA is >12. L2 proficiency is measured by the TOEIC test (Test of English in International Communication). Outcomes of the proficiency test are correlated by more / less years abroad (residence in an English-speaking country) and high / low self-rating of language skills of the high / low-proficiency L2 learner groups, respectively. Ojima and colleagues conduct ERP responses to visually presented experimental stimulus sentences, including semantic incongruities such as 'The house consists of ten *stories* / **cities* in total'. The behavioural

---

[17] In a similar study by Hahne and Friederici (2001), with native speakers and L2 learners of German (L1 = Japanese, AoA >10), the N400 as response to the auditory processing of semantic incongruity does not significantly differ between the native speakers and L2 learners.

[18] Hahne herself pointed out the large variation between subjects. She only reports the mean scores and standard deviations of the self-reported evaluation of German auditory comprehension, reading comprehension, and speaking and writing skills of her L2 learners. She does not give any range values of these scores.

results of the judgement task show that low-proficiency L2 learners perform less accurately in both conditions. Especially the incongruent sentences are rated more erroneously. Behavioural results between high-proficient L2 learners and monolinguals do not differ statistically. The ERP waves of all three groups show an enhanced N400 for the incongruent condition relative to the congruent condition (N400 effect). The authors report between-group differences in mean latencies and refer to proficiency as an indicator for the delay of the N400's peak latency: the more proficiency is improved, the shorter the delay (relative to monolingual processing). Ojima and his colleagues interpret their results against AoA influence. They put forward that quantitative differences in processing semantic incongruities are due to proficiency influences. They conclude that, with respect to timing, improvement of proficiency triggers more native-like processing, as indicated by the brain responses of L2 learners.

Moreno and Kutas (2005) carry out an ERP study on the processing of semantic incongruity (nouns and adjectives at the end of a sentence) with participants who are Spanish-English (L1-L2) bilinguals but vary with respect to language dominance: dominant / non-dominant language—e.g., Spanish / English and English / Spanish. The ERPs relative to the processing of incongruent words show an enhanced N400 component compared with ERPs associated with the processing of congruent words regardless of language dominance. Although either strength or distribution are determined by language dominance, the authors report latency delays of the N400 effect in the non-dominant language, irrespective of whether language Spanish or English. Accordingly, they carry out detailed correlation analyses including age of exposure, proficiency (conducted by the Boston Naming Task, henceforth BNT, letter and category fluency) and age with a subset of participants who all were exposed to their L1 at birth while to their L2 at a mean age (AoA) of 8. Results reveal correlations of all three factors with peak latency: increasing AoA and age determine a delay of the peak latency, whereas better proficiency indicates an earlier peak latency of the N400 effect. Further analyses including language dominance reveals a significant correlation with category fluency for the non-dominant language in that less fluency indicates delayed peak latency. Further, the authors report a marginal correlation of age of the participants and a latency delay of the N400 for the dominant language. Moreover, Moreno & Kutas conduct further correlation tests for both language dominant groups. Their results show that the delay of the N400 amplitude for the non-dominant language is mainly correlated with proficiency in the English / Spanish group whereas it is determined by AoA in the Spanish / English group. The authors acknowledge the inverse correlation between proficiency and AoA and further state that proficiency also determines the speed of processing a semantic incongruity in L1 (Spanish) whereas AoA does not (since it is acquired early). They conclude that the speed of processing a semantic incongruity is faster for the dominant language, and that both, AoA and proficiency, contribute to explain latency variation of the N400 in the non-dominant language. Thus, these findings indicate independent impacts on semantic processing by both factors, and contribute to unfold the correlation between AoA and proficiency influencing (L2) semantic processing and language dominance.

Newman et al. (2012) report behavioural and ERP evidence by English L2 learners (L1 = Spanish) and native English speakers when processing semantic incongruent words in sentences like 'The Irishmen sipped Todd's *whiskey* / **thunder* at the party'. The L2 learners all have an AoA >10. Proficiency is conducted for both L2 learners and native speakers by using samples of the TOAL test (Test of Adolescent and Adult Language). In their global analysis L2 learners are not grouped considering high / low proficiency, but proficiency is treated as a continuous variable. The behavioural results of the judgement task generally reveal better results for native speakers who equally do accurate rating for both types of sentences. L2 learners are less accurate and show a bias between conditions, which means that they are more erroneous in their ratings in cases of semantic incongruity. The ERP results reveal an N400 effect for the native speakers and L2 learners. The N400 effect appears stronger on the midline and frontal electrodes (irrespective of laterality) for native speakers relative to L2 learners. On the posterior electrodes, the N400 effect is equally strong for both groups, but it is delayed for the L2 learners. Newman and his colleagues further correlate proficiency scores with N400 amplitudes. They report that improvement of proficiency enhances the N400 effect especially on the left-lateral scalp-sites for native speakers and on a wider distribution for L2 learners. Additionally, L2 learners' ERPs demonstrate a latency shift, which, however, is not predicted by proficiency. The authors conclude that proficiency is an indicator of strength and lateralization of the N400 effect. Proficiency does not attribute to latency differences and possibly slower L2 (semantic) processing, which, then, might be correlated to AoA.

2.2.1.3 Summary and Prospects: Processing of Semantic Incongruity

In summary, all of the selected studies on L2 semantic processing reveal an enhanced N400 as response to semantic incongruity. Differences concerning the N400 effect between monolingual and L2 processing appear quantitatively —i.e., the N400 effect is delayed and sometimes reduced in L2 learners' ERPs, and additionally appears to be rather left than right-lateralized. All studies presume the integration view as the underlying functional role of enhanced N400 (see Chapter 2.2.1.1) and suggest that the quantitative differences are largely due to slower and less automatized processing with respect to L2 comprehension. In other words, L2 learners anticipate upcoming semantic information in their L2. The updating processes (i.e., integrating a word into the former context) appear more costly in L2 than in L1. The reasons for the enhanced processing costs in light of the AoA and proficiency levels of L2 learners remain unclear. As indicated above, results and interpretations are diverse. An open issue of most of these studies is the correlation between L2 learners' AoA and proficiency level. For instance, the results derived by Weber-Fox and Neville seem problematic since the groups with AoA <11 may not be viewed in the same L2 setting as the groups with later AoA. In the latter, English can truly be considered as L2, while Chinese is the mother tongue of all these participants. Therefore, for the group 1–3, it is debatable to attribute the results to the L1 / L2 distinction as they could also be accounted for language dominance. For those participants, Chinese seems to be the less dominant language and hence its status might be reconsidered with regard to language dominance rather than to the L1 / L2 distinction. It is less used and the mean results of the participants' self-ratings affirm that English has always been the language with which they feel more comfortable. A connected though not similar group imbalance concerning language dominance appears in the study of Moreno and Kutas. Language dominance is highly correlated with AoA and proficiency. However, the authors offer a detailed correlation analysis unfolding the correlation and stating that both factors make contributions to the differences between L1 and L2 semantic processing when language dominance is taken into consideration. The study by Ojima et al. controls for possible AoA effects, indicating that proficiency may override the latency delay and trigger more automatized semantic processing. Their data show a delay of the N400 effect elicited by the low-proficiency L2 group, which appears to be an AoA-related effect that may be compensated by improving proficiency, as is indicated by the data of the high-proficiency L2 group. Newman et al. also controls for AoA and further apply proficiency as an internal group factor, thereby revealing differences in the proficiency influence on semantic processing within different groups (i.e., monolinguals and L2 learners). It is different from the interpretation given by Ojima et al. as the delay of the N400 latency in L2 learners is attributed to AoA influence, while proficiency may impact the distribution and the strength of the N400 effect.

Despite the small differences of the N400 effects between L1 and L2 processing for the latter, it somehow appears that both AoA and proficiency claim influence over this. Against this background, the present study intends to investigate the questions (i) whether and (ii) how the influences impact the L2 processing of semantic incongruity? As to (i), the present study takes up the idea that L2 learners are sensitive to incoming semantic information but show enhanced processing costs relative to the processes of updating the previous context with the anticipated new information. The enhanced L2 processing costs have been revealed by differences of the N400's strength, distribution, and latency. The design of the present study is supposed to ensure any direct linking of both AoA and proficiency in terms of the strength, distribution, and latency of the potential N400 effect. According to question (ii), some of the above-reviewed data suggest a rather gradual than discontinuous influence of both factors. For instance, the results of the study by Newman et al study show that continuous improvement of proficiency gradually strengthens the N400 effect. The data of Ojima et al. suggests gradual development towards native-like latency of the N400 component (as response to incongruity) due to the improving L2 proficiency, albeit their categorical group design. Further, the data of Weber-Fox and Neville, who also group L2 learners, demonstrate a rather gradual latency change due to the increasing AoA. In order to provide valuable insights referring to the continuous character of influence, in the upcoming investigations both factors will be treated as continuous variables. Therefore, L2 learners will not be separated into groups on either AoA or proficiency bases. Chapter 4.2 comprises detailed descriptions of stimulus material, statistical models, hypotheses, and results for the processing of semantic incongruity by L1 speakers and L2 learners.

2.2.2 Double Nominative Violation

As briefly indicated in the beginning of this chapter, comparisons between ERP patterns as response to L1 and L2 syntactic processing reveal greater differences than those as response to semantic processing (for the latter, see Chapter 2.2.1.2). The structure that was chosen for investigation in the present study is the German double nominative violation (again, see example [3b]). On the basis of previous findings, brain responses of native German speakers usually reveal a biphasic N400-P600 processing pattern (see Chapter 2.2.2.2 below) when ERPs of double nominative violations are compared with the according ERPs as response to the processing of well-formed structures. Before turning to double nominative violations and how they are processed by L1 and L2 speakers, a short introduction is given to the P600 effect which is usually connected with the reflection of syntactic processing difficulties.

2.2.2.1 P600 and L1 Processing of Syntactic Violations

According to language processing, the P600 effect was first reported by Osterhout and Holcomb (1992). The authors investigated the processing differences of the so-called garden-path sentences, when temporal ambiguity is resolved either by the preferred or non-preferred reading as in 'the broker ?hoped / persuaded to sell the stock *was* sent to jail'[19], respectively. In case of the non-preferred reading, the P600 component appears to be enhanced in comparison with the preferred reading, thereby resulting in a P600 effect. Physically, the P600 component is a positive-going wave, usually starting 600 ms after stimulus presentation (critical item). The P600 effect typically has a posterior distribution— i.e., it appears largest on the posterior scalp-regions.[20] Osterhout & Holcomb account for their finding of the P600 with higher loads of processing, since simple processing principles have to be overruled and the sentence's former reading has to be reanalysed. Kaan, Harris, Gibson, and Holcomb (2000) find that L1 processing of syntactically more complex sentences produces an enhanced P600 when compared with the processing of simpler structures. The authors maintain that whenever a structural integration problem of an item into the preceding sentential context occurs, there is an increase in demands with regard to the language processing system. Next to the enhanced processing demands due to the increased syntactic complexity and the resulting reanalysis of a syntactic structure, the P600 also occurs with a structural violation (ungrammaticality) that has to be repaired (see Friederici, 1995, 2002). Principally, the P600 has been observed as response to phrase structure and sub-categorization violations (again, see Kaan et al., 2000; see also Hahne & Friederici, 1999; again see Osterhout & Holcomb, 1992), as well as verb-argument-agreement violations concerning number, tense, gender, and case (e.g., Barber & Carreiras, 2005; Bott, 2009; Frisch & Schlesewsky, 2001; Gunter, Friederici, & Schriefers, 2000; Hagoort, Brown, & Groothusen, 1993). The P600 elicited by a syntactic violation often follows early ERP components associated with automatic structure-building processes like left-anterior negativities (LAN) (e.g., Barber & Carreiras, 2005; Coulson, King, & Kutas, 1998; Gunter et al., 2000; Muente, Heinze, & Mangun, 1993[21]; Osterhout & Mobley, 1995).

---

[19] '?'-indexed item indicates the non-preferred reading.

[20] The term 'P600' will be used throughout the remainder of this thesis, when referring to the 'P600 effect'.

[21] Muente et al. (1993) find a broadly distributed early negativity with a frontal maximum as response to grammatical violation of word pairs (i.e., agreement mismatch of personal pronoun + verb and

Accordingly, the P600 has been interpreted as the reflection of a succeeding process of repairing a syntactically ill-formed structure. For an extensive review on the P600, the reader is referred to Gouvea, Phillips, Kazanina, and Poeppel (2010).

According to the mechanisms of syntactic repair, the P600 also occurs subsequent to an N400 effect. Frisch and Schlesewsky (2001, 2005) report a P600 following an N400 effect when a structure shows identical case-marking on both NP arguments within one clause (in German). They interpret the P600 as a reflex of syntactic repair mechanisms due to the fact that identical case-marking on two arguments of the same verb is ungrammatical (see below for a detailed review of ERP studies on the processing of German double case-marking violations). A syntactic interpretation of the P600, following an N400 effect, is also given by studies concerning the processing of non-licensed German NPI 'jemals' ('ever') (e.g., Drenhaus, Saddy, & Frisch, 2005). Here, the underlying assumption is that syntactic linearity is important for appropriately processing the NPI constructions. A P600 occurs when the proper licensor (e.g., negation) is missing or structurally inaccessible to the upcoming NPI (see also Chapter 2.2.3.1 for a more precise look on ERP evidence revealed by non-licensed NPIs).

The shortly reviewed findings of the P600 indicate that its functional role entails processing mechanisms related to structural reanalysis and / or repair. Conceivably, whenever the language processor encounters a problem due to structural violation or enhanced syntactic complexity, neural mechanisms reflecting processes to resolve the problem / failure are activated.[22] Additionally, the P600 appears enhanced when a grammatical judgement task is enforced, which directly triggers the processes of repair

---

possessive pronoun + noun). However, they do not report a subsequent positivity effect (i.e., P600) reflecting syntactic reanalysis/repair, which is possibly due to the missing sentence context.

[22] More recently, the P600 has also been demonstrated as response to the processing of semantic reversal anomalies (van Herten, Kolk, & Chwilla, 2005) and semantic verb-argument violations (Kuperberg, 2007). In line with this, the P600 is assumed to reflect a more general component of conflict monitoring (van Herten, Chwilla, & Kolk, 2006). Additionally, Brouwer and his colleagues (2012) argue that instead of being purely syntactic characteristics, the P600 resembles processes of integration due to the updating of the mental representation with new information. Therefore, it is enhanced and reflects difficulties to interpret the whole structure. Notwithstanding this, for the purpose of the present thesis and the investigations on syntactic processing, the expectation of the occurrence of the P600 as a reflection of the costly processing devoted to the retrieval of mechanisms of syntactic reanalysis and repair (see above) shall be sufficient.

and reanalysis (Hahne & Friederici, 2002). Differences in strength, latency, and topographical distribution are less discussed than for the N400 effect (see Chapter 2.2.1.2 above). Gouvea et al. (2010) propose that retrieval mechanisms due to syntactic complexity and ambiguity may determine the latency of the P600, while mechanisms concerning phrase structure building processes are reflected by its strength. For the present thesis it shall be sufficient to link the P600 to syntactic repair mechanisms. Furthermore, the strength, rather than latency and distribution, of the P600 will be of main interest with respect to the L2 processing of syntactic violations and the potential differences between L1 and L2 processing. As will be seen below, the P600 has also been elicited as response to syntactic L2 processing. The underlying differences between L2 learners' and L1 speakers' ERPs have been associated with both AoA and proficiency influence. They will be returned to in more detail in Chapter. 2.2.2.3, below.

2.2.2.2 N400-P600 and L1 Processing of German Double Nominative Violation

In his dissertation, Frisch (2000, Experiments 5 and 6) studies the processing of double nominative violation in German. In Experiment 5, the stimulus material has an NP V NP word order[23], as in 'Welcher$_{\mathbf{NOM}}$ Kommissar$_{\mathbf{NOM}}$ lobte *den*$_{ACC}$ *Detektiv*$_{ACC}$ / **der*$_{NOM}$ *Detektiv*$_{NOM}$ im Radio' ('Which$_{\mathbf{NOM}}$ detective$_{\mathbf{NOM}}$ admired *the*$_{ACC\ /\ *NOM}$ *agent*$_{ACC\ /\ *NOM}$ in the radio'). In Experiment 6, the word order is NP NP V (see Footnote 23), as in 'Hans weiß, welcher$_{\mathbf{NOM}}$ Kommissar$_{\mathbf{NOM}}$ *den*$_{ACC}$ / *der*$_{*NOM}$ *Detektiv*$_{ACC\ /\ *NOM}$ gelobt hat' ('Hans knows, which$_{\mathbf{NOM}}$ detective$_{\mathbf{NOM}}$ *the*$_{ACC\ /\ *NOM}$ *agent*$_{ACC\ /\ *NOM}$ admired [has]').[24] ERP results of both experiments reveal a biphasic N400-P600 processing pattern, when the

---

[23] In German, SVO is the linear word order in matrix sentences. Generally, it is assumed that the underlying German word order is SOV. In order to form a matrix clause, the verb has to be moved to the second position. This restriction is also called V2 word order (see e.g., Fagan, 2009 for a coherent introduction to the German word order).

[24] Frisch (2000, Experiments 5 and 6) also investigates the ill-formed double accusative violations in German. Results show an enhanced N400 relative to the N400 evoked by double nominative violations, when the verb precedes the second NP. There are no differences when the verb occurs after the second NP.

second NP is encountered as nominative. Frisch interprets these results as—independent of verb position and therefore availability of verbal thematic information—identical case-marked NPs enter into a competition on the same structural position and thematic interpretation. The N400 effect is evoked by the case features of the arguments (cf. Frisch, 2000, p. 235)—i.e., as soon as the processor encounters the second DP (also being marked nominative), the structure is not interpretable anymore.[25] The subsequent P600 infers enhanced processing costs reflecting the enhanced repair mechanisms to resolve the structural competition on position and, hence, meets a structure that is interpretable (ibid, p. 250).

Frisch and Schlesewsky (2001) investigate the processing on double nominative violations within a similar NP NP V structure, as in Frisch (2000, Experiment 6), thereby further alternating the animacy of the second NP. They report the occurrence of an N400-P600 processing pattern only when both NPs are identically case-marked and both are animate. In case of an inanimate second NP, only the P600 occurs and the N400 effect stays absent. Concerning the P600, the authors suggest that it indexes repair mechanisms due to the ungrammatical structure (see also above), which is also observable when both NPs have identical case-marking but different animacy marking. Their interpretation of the occurrence / absence of the N400 suggest that identical case-marked and animate arguments cannot easily be thematically hierarchized due to animacy information. They argue that the language processor has no possibility to overrule the case-marking violation by thematically hierarchizing the subject and object NPs due to animacy (see also Frisch & Schlesewsky, 2005; Frenzel, Schlesewsky, & Bornkessel-Schlesewsky, 2011). In other words, the N400 is enhanced in contexts where thematic hierarchizing is challenged because animacy information does not resolve the processing problems due to ill case-markings. This extends the argumentation of Frisch (2000, see above), where the enhanced N400 is associated with

---

[25] In a study by Coulson et al. (1998), a LAN-P600 pattern occurred in native English speakers when processing a case-violation on the personal pronoun, as in sentences like 'the plane took *us* / **we* to paradise and back' (critical item is in italics, *-indexed item indicates case-marking failure). Their interpretation of the occurrence of LAN, along with the enhanced working memory demands, refers to the search for an adequate referent in the prior discourse of the sentence.

costly processing because of the case features of the critical item. This also suggests that there are differences between the sources of neural mechanisms retrieved for the processing of semantic incongruity (as described in Chapter 2.2.1.1) and for thematic processing. Still, the latter also accounts for an integrative view as thematic processing involves the assignment of thematic roles, which is understood as also being part of updating the previous context. If a thematic role cannot easily be assigned to an upcoming NP, integration of that NP into the former sentence context is constrained and hence processing difficulties appear.

2.2.2.3 N400-P600 and L2 Processing of German Double Nominative Violation

The majority of studies on L2 processing of German double nominative violation are based on the processing of a so-called artificial language, namely Mini-Nihongo, a miniature version of Japanese (for a coherent description of Mini-Nihongo, see Mueller, 2005; see also Mueller, Hahne, Fujii, & Friederici, 2005). Mueller et al. (2005) record the accuracy judgments and ERPs by Japanese native speakers,[26] and L2 learners of Mini-Nihongo with AoA >18 and German as their L1. L2 learners are grouped on the basis of their L2 proficiency—i.e., one group was trained in Mini-Nihongo, while the other was not. Mueller and colleagues conduct their data as response to the processing of double nominative violations like 'Ichi wa no kamo-ga$_{\mathbf{NOM}}$ ni hiki no *neko-o*$_{ACC}$ / *-*ga*$_{NOM}$ tobikoeru tokoro desu' ('One duck$_{\mathbf{NOM}}$ jumps over two *cats*$_{ACC\ /\ *NOM}$'). Results reveal a biphasic N400-P600 pattern for native Japanese speakers.[27] They interpret this ERP pattern in the light of Frisch's and Schlesewsky's finding (see Chapter 2.2.2.2

[26] Mueller et al. (2005) say that native Japanese speakers are considered as native Mini-Nihongo speakers since the lexical items and grammatical rules are part of Japanese and therefore resemble a valid model of native Japanese.

[27] The N400 as response to the case-violation starts very early and shows two peaks especially on the frontal electrodes. Mueller et al. (2005) suggest that two distinct components might be considered here. Next to the thematic N400, they relate the earlier negativity component to some reflection of the mismatch of phoneme expectation. It means that a semantic context triggers the expectation of a certain phoneme. If that expectation is not met, a very early frontally distributed negativity is revealed.

above). The enhanced N400 as response to the case-violation reflects difficulties to thematically hierarchize two identically case-marked and animate NPs. The subsequent P600 signals mechanisms of repair of the ill-formed structure. With respect to the same N400 and P600 time-windows, the ERPs of untrained L2 learners do not show any effect at all. Additionally, the behavioural results (i.e., the accuracy of grammatical judgements) of untrained L2 learners concentrate far below chance level. ERPs of the trained L2 learners reveal a robust P600 but tend to lack the occurrence of the N400 effect. The authors interpret this absence of the N400 effect twofold: It either reflects the impairment of thematic processing or a more general lack of the activation of processing resources, which then reveals some kind of floor effect due to the observation that the ERPs as response to the case-congruent condition are more negative-going compared with those elicited by native Japanese speakers. Both approaches are accounted for AoA influence. The authors further draw attention to the differences between the observed N400 effects in studies yielding the L2 processing of semantic incongruity. They note that the contextual integration of the actual meaning of a specific word might be less difficult to be acquired in the course of L2 learning than thematic processing and the entailed hierarchizing of semantic roles (cf. Mueller et al., 2005, p. 1240). Mueller et al. further suggest that the canonicity of the stimulus material might have triggered a purely syntactic processing strategy that blocks any thematic processing mechanisms. They relate this idea to the observation that although L2 learners were trained in Mini-Nihongo, accuracy ratings differed between native speakers and trained L2 learners, as the native speakers reached significantly higher accuracy. The P600 elicited by the trained L2 learners resembles that of native Japanese speakers. The authors conclude that L2 learners are able to activate the neural mechanisms associated with syntactic repair and triggered by the ungrammaticality of the relevant structure. They further infer the occurrence of the P600 to be impacted by L2 proficiency, as the untrained group with the same AoA as the trained group did not reveal any late positivity effect.

In a following study, Mueller, Hirotani, and Friederici (2007) carry out an ERP study by using the same double nominative violations in Mini-Nihongo with native Japanese speakers and highly trained L2 learners (AoA >18, L1 = German).[28] In this study, the authors tried to ensure that both groups do not differ with respect to Mini Nihongo proficiency, which improves the experimental design of their study in 2005 (reported above). The authors do not find any significant differences in the accuracy rating scores between the two groups. Their ERP results, again, show a biphasic N400-P600 pattern for the violation condition in native Japanese speakers. L2 learners show a LAN-P600 pattern as response to the double nominative violation. With regard to the P600 responses, no differences are reported between native speakers and L2 learners. This finding resembles the outcome of their former study (Mueller et al., 2005), where the P600 elicited by trained L2 learners did not differ from the P600 responses of native speakers. The occurrence of LAN in L2 learners' ERPs is different from their former results (Mueller et al., 2005). First of all, a LAN is not observed in the earlier results. Second, this processing pattern differs from the one observed for native Japanese speakers. The authors interpret the LAN-P600 pattern by saying that L2 learners use a simplified processing strategy with restriction to formal aspects, and also that L2 learners rely on the fact that an occurrence of a certain case-marker on the first NP requires the occurrence of another case-marker on the next NP. Thus, there is evidence of mere syntactic L2 processing. However, the occurrence of LAN in only highly-trained L2 learners indicates proficiency influence. Hence, the improving proficiency yields refined structural processing mechanisms in L2. The lack of the N400 effect underlines the assumption that thematic processing mechanisms are not activated in L2 processing, inferring AoA influence. The study by Mueller, Girgsdies, and Friederici (2008) further investigate on whether this lack of N400 may truly be accounted for AoA influence. The author acknowledges that especially in the study by Mueller et al. (2007) concerning L2 learners' training in Mini-Nihongo, major emphasis has been given to the correct semantic interpretation due to word order and morpho-syntactic rules, which might have blocked the activation of neural resources linked to thematic processing mechanisms from the beginning. Therefore, Mueller et al. (2008) investigate

[28] Additionally, Mueller et al. (2007) study the processing of double accusative violations. Since the present thesis does not focus on such structures, the results will not be reported.

the L2 processing (L1 = German, AoA >18) of the same structures, as in Mueller et al. (2005, 2007). They change the L2 learners' training modality of Mini-Nihongo by completely removing the semantic load. The authors argue that, in this case, the activation of integration mechanisms is facilitated and that, consequently, the retrieval of thematic hierarchizing might not be blocked. Brain responses of L2 learners reveal an N400-P600 pattern with respect to the processing of the case-violation. However, Mueller and her colleagues do not compare the L2 learners' ERP pattern to any of their formerly conducted native speakers' ERP pattern (cf. results by Mueller et al., 2005, 2007). Thus, there are limits to the argument that L2 learners show an ERP pattern that differs from that of formerly conducted L2 learners but not from native Japanese speakers. Further, my own visual inspections of the data by Mueller et al. (2007) and Mueller et al. (2008) suggest that both early ERP effects are similar with respect to strength and distribution. Their statistical analysis of both effects reveals enhanced differences between syntactic conditions on anterior and left-lateral scalp sites. Therefore, the negativity effect elicited by Mueller et al. (2008) may also be related to a LAN-like negativity, rather than an N400-like negativity. Furthermore, this would support the idea that highly proficient L2 learners are able to activate the neural resources of refined syntactic processing mechanisms. However, they do not show any sensitivity to activate those neural resources associated with thematic processing for case-violation because the latter may then be accounted for AoA influence.

In a further Mini-Nihongo study by Mueller (2009), she investigates the L2 processing of double-nominative violations with either familiar or unfamiliar vocabulary (nouns). Again, L2 learners have German as L1, AoA >18, and are trained to high L2 proficiency. Her results show an N400-P600 processing pattern for L2 learners when they are lexically familiar with the critical noun (case violation). The N400 effect stays absent when the critical noun is lexically not familiar to the L2 learners. Her interpretation indicates that the N400 effect is a reflection of problems with thematically hierarchizing NP arguments (as indicated by Frisch & Schlesewsky, 2001, see above). The P600 reflects syntactic repair mechanisms. Interestingly, the proficiency of the L2 learners is the same as in the study by Mueller et al. (2007). It is therefore questionable whether the occurrence of the N400 effect is truly influenced by proficiency. I tentatively suggest that the occurrence of the N400 effect with familiar words (wrongly case-marked) might be due to the experimental design which here triggers lexical processing mechanisms via the familiarity condition rather than syntactic processing mechanisms, as has been demonstrated by the former studies of Mueller et al. (2005, 2007, and also 2008).

Domke (2012, see also appendix 6) investigates the syntactic L2 processing of German double nominative violations and reports the processing of differences between L2 learners of German (L1 = Polish) with different AoA. In her study, all L2 learners are equally highly L2 proficient but differ in terms of their AoA. Domke separates the L2 learners into four groups with regard to the varying AoA, namely 0–3, 4–6, 6–8, and 8–10. The processing of case-marking failure is tested with German matrix sentences like 'Der$_{\mathbf{NOM}}$ Mann$_{\mathbf{NOM}}$ pflanzt *den*$_{ACC}$ *Baum*$_{ACC}$ / **der*$_{NOM}$ *Baum*$_{NOM}$ …' ('The$_{\mathbf{NOM}}$ man$_{\mathbf{NOM}}$ plants *the*$_{ACC\ /\ *NOM}$ *tree*$_{ACC\ /\ *NOM}$ …'). The critical NP in her material is non-animate; therefore, demands with regard to the activation of underlying neural resources related to the processing mechanisms of thematic hierarchizing should be alleviated. Her results show a LAN-P600 pattern only in ERPs by L2 learners with AoA 0–3. L2 learners who started L2 acquisition later than at the age of three years only elicit a P600 that additionally is attenuating as AoA increases. She interprets the results more in the line of higher processing demands on morpho-syntactic issues, rather than thematic integration problems. This, in part, resembles the results and interpretations of Mueller et al. (2005, 2007, see above). However, the findings concerning the P600 amplitudes are different, as they decrease with growing AoA, irrespective of equal high levels of L2 proficiency. Domke argues that not only early syntactic processing but also the activation of the controlled mechanisms of syntactic repair seem to be sensitive to AoA influence. Yet, the low number of participants seems problematic for this study. Also, non-availability of data from a monolingual control group, which would have raised validity with respect to interpretations towards differences between syntactic L1 and L2 processing mechanisms, is another concern.

2.2.2.4 Summary and Prospects: Processing of Double Nominative Violation

The results of the above-reviewed L1 and L2 ERP data as response to the processing of double nominative violations indicate the following: (i) ERPs of double nominative violations yield brain responses that reflect the activation of neural resources associated with thematic and syntactic processing; (ii) the ERP patterns elicited by L1 and L2 processing differentiate; and (iii) both factors, the L2 learners' AoA and proficiency, influence the processing of a double nominative violation and hence cause differences with the L1-processing patterns.

With respect to (iii), the underlying ERP data appears different from that associated with the processing of a semantic incongruity (again, for the latter see Chapter 2.2.1.2). Studies on the processing of double nominative violations in Mini-Nihongo strongly controlled for the correlation between AoA and proficiency: AoA always was >18 and proficiency was subject to experimental manipulations. To be more precise: When AoA is held constant, improvement of the L2 proficiency seems to highly impact L2 syntactic processing. Moreover, this influence is different for early ERP components, such as LAN, and late ERP patterns, such as P600. Generally, the impact of improving proficiency on both brain responses seems to be continuous. However, and most interestingly, proficiency influence only becomes visible given a certain threshold is achieved, and the limitation for a threshold varies with respect to the two brain responses, which suggests that L2 proficiency itself sets the threshold limitations. Ponderably, while LAN-like negativities only occur at very high levels of L2 proficiency, the P600 effects appear rather native-like at moderate or intermediate L2 proficiency levels already. Against the background of the Mini-Nihongo data reported above, Steinhauer, White, and Drury (2009) propose a model of temporal dynamics, indicating that improvement of proficiency triggers the native-like syntactic processing. They identify six stages of grammatical processing reflected by ERP patterns, which step-by-step approach native-likeness.[29] In this respect and with regard to the activation of neural mechanisms inferring syntactic processing (i.e. P600), Steinhauer et al. (2009) propose that improving proficiency may compensate for AoA influence as the latter is only clearly visible at low L2 proficiency levels (again see the results of Mueller et al., 2005). Nevertheless, as could be seen above, AoA impact is visible for the processing of double nominative violations even for highly proficient L2 learners. This impact appears in the absence of an early negativity effect (which, in contrast, is reported for all related L1-processing patterns). Although improvement of proficiency may trigger the occurrence of an early negativity, it does not resemble a classic N400 effect.

To sum up, on the basis of the above-reviewed data, there seems to be an interactive influence of AoA and proficiency on the activation of neural mechanisms responsible for the mere syntactic L2 processing of double nominative violations. Furthermore, there seems to be rather clear AoA influence on the activation of neural resources associated with the thematic processing mechanisms in L2. The present study intends to predominantly investigate the impact of the interaction between both factors on syntactic L2 processing. There are two main objectives (similar to those proposed for semantic L2 processing as described in Chapter 2.2.1.3): (i) whether and (ii) how AoA and proficiency impact the L2 processing of double nominative violations? As to (i), the potential differences between L1 and L2 ERPs will be of prime concern. With respect to (ii), it will be emphasized whether changes in syntactic L2 processing are truly subject to a proficiency threshold—i.e., whether a certain level of proficiency has to be reached to reliably activate neural mechanisms of syntactic repair or whether proficiency influence, in general, appears rather gradually. Moreover, the present study aims to investigate whether the potential interaction between AoA and proficiency (i.e., improving proficiency compensates for AoA influence) is statistically observable in the upcoming data. Therefore, the present study will strictly rely on syntactic processing

[29] Corresponding proficiency influence on the development of L2 learners' ERPs (as response to syntactic violation processing) towards native-like patterns is also reported in longitudinal studies; see McLaughlin et al. (2010), and Osterhout et al. (2006), and more recently Tanner et al. (2013). However, the results of these studies predominantly reflect brain responses to the processing of verb-argument violations, not double case violations.

and not on potential processes underlying thematic hierarchizing. In other words, the activation of neural resources associated with thematic hierarchizing should not be enhanced in the present study to clearly reveal syntactic processing patterns (here P600). To ensure this, the animacy status of the critical item is not modulated and all critical items are non-animate. See Chapter. 4.3 for the stimulus material, statistical models, hypothesis, and the behavioural and ERP results of the processing of double nominative violation, as revealed by L1 speakers and L2 learners.

2.2.3 Negative Polarity Items (NPIs)

The monolingual processing of the non-licensed German NPI 'jemals' ('ever') involves the activation and retrieval of both semantic and syntactic neural mechanisms. Usually, a biphasic N400-P600 pattern is revealed when ERPs as response to the processing of the licensed NPI are compared with ERPs elicited by the processing of the non-licensed NPI. To my best knowledge, it is yet to be investigated whether analogous L2 ERP patterns (as response to licensed vs. non-licensed German 'jemals' ['ever']) appear differently, and whether AoA and / or proficiency claim influence. Against the background of the findings based on semantic and syntactic L2 processing patterns, as described in Chapters 2.2.1 and 2.2.2, respectively, there might be some hypotheses drawn with regard to potential differences between L1 and L2 ERP patterns revealed by the processing of non-licensed NPI structures.

2.2.3.1 N400-P600 and L1 Processing of Negative Polarity Items

There has been quite intense research on the L1 processing of the German NPI 'jemals' ('ever'). Saddy, Drenhaus, and Frisch (2004) compare the processing of NPI 'jemals' ('ever') in an appropriate licensing context '**Kein Mann**, der einen Bart hatte, war *jemals* froh' ('**no man** who had a beard was *ever* happy')[30] with an inadequate licensing context—i.e., when there is no licensor, as in '**Ein Mann**, der einen Bart hatte, war **jemals* froh' ('**a man** who had a beard was **ever* happy')[31]. Their results show an enhanced N400 component for the NPI '*jemals*' ('ever') in the non-licensed context. Saddy et al. interpret this N400 effect as a reflex of semantic integration problems. It means that as soon as the language processor encounters the NPI in a non-licensed context, the problems to integrate the NPI emerge. In a following study, Drenhaus et al. (2005) investigate the processing of the same structures, as in Saddy et al. (2004), and add one structure that contains a negation which, however, is structurally not accessible and therefore renders the sentence ungrammatical, as in 'Ein Mann, der **keinen Bart** hatte, war **jemals* froh' ('a man who had **no beard** was **ever* happy') [32]. Their results again reveal an N400 effect that additionally is modulated by violation. The condition where the negation is inaccessible but still appears prior to the NPI elicits reduced N400 amplitudes in comparison with the condition without any prior negation. This modulation of the N400 effect is associated with enhanced semantic processing costs that are apparently lowered when a licensor occurs, although it is structurally not accessible. The authors also report a subsequent P600 that was not present in their former results (i.e., Saddy et al., 2004).[33] The P600 is not modulated by violation but equally enhanced for both violation conditions. Drenhaus and his colleagues interpret these results as evidence that NPI licensing not only depends on semantic but also on syntactic conditions. The syntactic processing problems are reflected by an enhanced P600 component and indicate the activation of neural resources associated with syntactic reanalysis mechanisms (see Chapter 2.2.2.1). These mechanisms are activated because the NPI does not occur within the scope of the appropriate licensor due to absence and / or inaccessibility of such. According to the strength of an appropriate NPI licensor, Drenhaus, Blaszczak, and Schuette (2007) investigate the processing of

[30] Licensor (negation) is bold-indexed.
[31] Inappropriate licensor (no negation) is bold-indexed.
[32] Inaccessible licensor (negation) is bold-indexed.
[33] An SRA analysis (see Footnote 6) on the data by Saddy et al. (2004) reveals a subsequent P600 effect, which is interpreted in line with enhanced syntactic processing mechanisms (Drenhaus, beim Graben, Saddy, & Frisch, 2006).

NPI 'jemals' ('ever') in different licensing contexts, namely negation and wh-context[34], in comparison with two non-licensed definite and indefinite contexts ('**Kein / Welcher / *Ein / *Der** Lehrer hat den Schüler *jemals* geschlagen?' ('**No / Which / *A / *The** teacher has the student *ever* hurt?')[35], see also Drenhaus, Blaszczak, & Domke (under review) and Schuette (2006). Their overall results show a biphasic N400-P600 pattern of the two non-licensed contexts in comparison with the negation context. Furthermore, the N400 is modulated. ERPs revealed by the NPI in the indefinite non-licensed context show weaker N400 amplitudes than ERPs as response to the definite non-licensed context. In comparison with the wh-context, ERPs relative to the NPI reveal an enhanced N400 only in the definite non-licensed condition. On the basis of their results the authors put forward that the language processor is sensitive towards the NPI licensing strength.[36] Moreover, the integration of an NPI into a wh-context seems less probable than into a negative context, which is revealed by the modulation of the N400. Differences in the P600 in relation to licensing strength have not been found. The P600 occurs robustly in all non-licensed vs. licensed comparisons.

In a more recent study, Yurchenko et al. (2013) examine the processing of the Dutch NPI 'bijster' ('at all') in affirmative and negative contexts such as '…en / maar zeijn handschrift was **neit / *ook** *bijster* leesbaar' ('and / but his handwriting was **not / *also** *at all* readable')[37]. Their results show an enhanced N400 whenever the NPI occurs in an affirmative rather than in a negative context. This indicates processing difficulties due to the absence of an appropriate licensor and partly corresponds to the findings concerning the processing of non-licensed German NPI 'jemals' (see above). The authors do not report any subsequent positivity effects (as was reported for the processing of non-licensed 'jemals' ('ever'), see above). Yurchenko et al. account for the

[34] The wh-context refers to the umbrella term wh-question that includes questions formed with an interrogative word. In English, these interrogatives preponderantly start with wh (e.g., what, which, who, etc.).
[35] Appropriate and inappropriate licensors (negation, wh-element, indefinite, and definite determiner) are bold-indexed.
[36] The NPI licensing strength refers to the hierarchy of contexts that are proper licensors of NPIs. This has been suggested by van der Wouden (1997).
[37] Appropriate and inappropriate licensors (negation, no negation) are bold-indexed.

absence of the P600 in that syntactic processing costs are not enhanced due to the pure semantic violation, as the licensor is simply not there. They further point out that the structure of the experimental stimuli used in studies with German NPI 'jemals' is more complex due to the longer distance between the licensor and the NPI. Yurchenko et al. suggest that the processing of such complex structures is generally more difficult. Thus, in their view, only this difficulty—i.e., due to structural complexity—is indexed by the increased P600 (for related comments see also Steinhauer, Drury, Portner, Walenski, & Ullman, 2010). However, the sentence structure in the studies by Drenhaus et al. (2007, under review) is not as complex as in the earlier study by Drenhaus et al. (2005). Hence, the complexity argument based on long-distance dependencies appears to be limited. Yurchenko et al. also note that the enhancement of the P600 might be triggered due to the grammatical judgement task prompted by the studies with NPI 'jemals', where the difficulty regarding decisions is enhanced *per se* (see again Hahne & Friederici, 2002; see also Chapter 2.2.2.1).

2.2.3.2 N400-P600 and L2 Processing of Negative Polarity Items

As to my best knowledge, there has so far been no published ERP data on NPI processing elicited by L2 learners. Intensive research in the World Wide Web offered one set of ERP data on NPI processing by L2 learners of Basque (L1 = Spanish). Pablos, et al. (2011, informally published data cited with permission of L. Pablos, personal communication December 12, 2013) report differences between the L1 and L2 processing of Basque NPI 'ezer' ('anything'). All L2 learners are highly proficient in Basque and report an AoA between 2 and 4. The stimulus material consists of different licensing violations for object Basque NPI 'ezer' ('anything'), namely absence of negation as in 'Gizonak **ba**darama garrantzizko **ezer* poltsikoan.' ('Man.the does.carries importance.of **anything* pocket.the.in'), negation inaccessibility as in 'Gizonak darama **ez** garrantzizko **ezer* poltsikoan' ('Man.the carries not importance.of *anything* pocket.the.in')[38] and scope violation such as in 'Gizonak **ez** du **eraman** garrantzizko ?*ezer* poltsikoan' ('Man.the **not** aux carry importance.of ?*anything* pocket.the.in')[39].[40] Behavioural results by native speakers of Basque reveal differences in the acceptability rates for the non-licensing contexts so that the absence of negation is rated 57% acceptable and the scope violation is 81%. A clear acceptability rejection is only found when the licensor is inaccessible (95%). ERP results by L1 speakers support the results of the acceptability ratings: only the inaccessibility condition, which is rated highly inacceptable, reveals a reliable N400 effect (relative to control condition, see Footnote 40). The researchers assume that the occurrence of the N400 effect reflects the activation of the processing mechanisms of semantic and syntactic anomaly due to the inaccessibility of the appropriate licensor for the NPI. Subsequent P600 responses are not reported. They interpret the absence of the P600 in terms of the enhanced lexical complexity of the object NPI that reveals stronger influences on the processing of semantic congruity than adverbial NPIs like German 'jemals' 'ever' (see above). For L2 learners of Basque, the behavioural ratings are similar to those revealed by native speakers in all conditions. Again, only the inaccessibility condition was clearly considered inacceptable (83%). However, ERPs of L2 learners do not reveal an N400 effect relative to the control condition. A subsequent P600 is not reported, either. The authors claim that L2 processing is not (very) sensitive to the syntax-semantics interface; hence, unlike L1 processing, neural mechanisms are not activated in response to inappropriate NPI licensing. Since all L2 learners were highly proficient and all had early AoA, it is rather fussy to account for a direct influence of either one of the factors or their correlation.

[38] Inaccessible licensor 'ez' ('not') is bold-indexed.
[39] Scope violating licensor 'ez' ('not') is bold-indexed '
[40] ERPs of violation conditions were compared with those of a correct control condition: According to the above-mentioned examples, the corresponding control sentence was 'Gizonak **ez** darama garrantzizko *ezer* poltsikoan' ('Man.the **not** carries importance.of *anything* pocket'). Licensor is bold-indexed.

2.2.3.3 Summary and Prospects: Processing of Negative Polarity Items

In summary, the processing of German non-licensed NPI 'jemals' ('ever') reveals an N400-P600 ERP response by native speakers. The N400 effect is interpreted in the light of the integrative account and hence associated with the enhanced integration difficulties (see Chapter 2.2.1). It means that the former context cannot easily be updated with the upcoming NPI since there is no licensor available. The P600 is supposed to reflect the enhanced mechanisms of syntactic reanalysis / repair since the upcoming non-licensed NPI renders the structure ungrammatical. With regard to differences between L1 and L2 processing of NPI constructions, the scarce data suggests that integration difficulties of a non-licensed NPI are not enhanced in L2 processing, even at very early AoA and high L2 proficiency.

With respect to the correlation of AoA and proficiency and its influence on L2 processing, the present study investigates the processing of licensed and non-licensed German NPI 'jemals'. The processing of such structures enforces the activation of semantic and syntactic processing mechanisms. The following four objectives will be central: (i) whether there are differences between L1 and L2 processing patterns; (ii) whether there is clear dominance for the activation of either semantic or syntactic L2 processing mechanisms; (iii) whether L2 processing is influenced by AoA and / or proficiency; and (iv) whether the potential AoA and proficiency influences are similar in accordance with those predicted for isolated L2 semantic (see Chapter 2.2.1.3) and syntactic processing (see Chapter 2.2.2.4). Materials, hypotheses, and results are given in Chapter 4.4.

2.3 A Model of L2 Processing

The previous review of ERP studies on L1 and L2 semantic and syntactic processing indicate that L2 semantic processing patterns are less different from L1 processing patterns than are L2 syntactic processing patterns. Up to date, within SLA research, there is no sufficient answer to why some linguistic phenomena (predominantly syntactic ones) produce differences between L1 and L2 processing and others (e.g. lexical-semantically related ones) do not.[41] One promising approach that may contribute to the discrepancy between lexical-semantic and syntactic processing and the above mentioned differences between L1 and L2 processing is the dual mechanism approach to language processing suggested by Clahsen & Felser (2006a, b). The authors postulate the "shallow structure hypothesis". This account attributes to the dissociation between semantic and syntactic processing. Likewise, the shallow structure hypothesis is based on the idea that for successful sentence comprehension the processing system has to follow a lexical-semantic processing route on the one hand and a grammatical-syntactic processing route on the other hand.[42] The former is based on strategies that are retrieved by lexical-semantic processing heuristics. The latter involves mechanisms of full grammatical parsing and structural analysis.

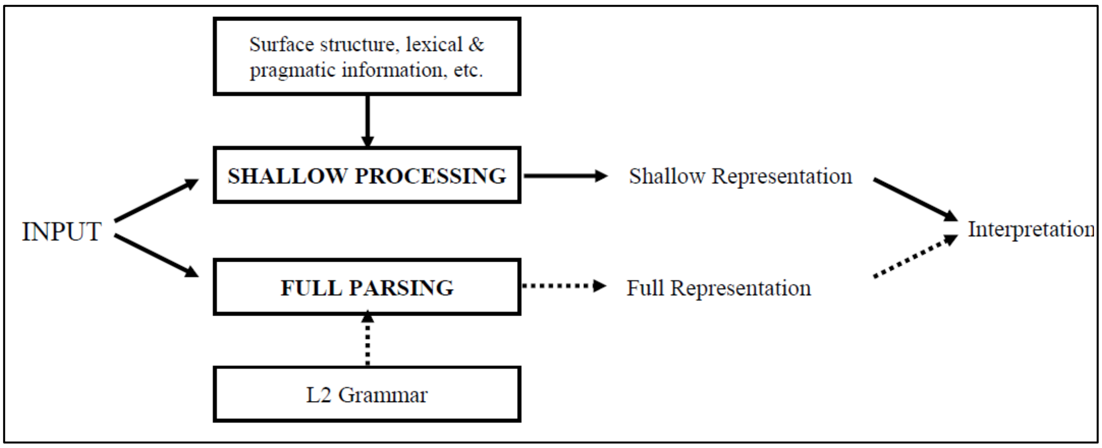


Figure 2.3: Adopted from Clahsen & Felser (2006b, p.119) illustrating the model of L2 processing according to the shallow structure hypothesis. Original citation: "Figure 1. Of the two routes to interpretation available in principle, full parsing is restricted in L2 sentence processing because of inadequacies of the L2 grammar."

41 Within SLA research there is the long enduring debate whether L1 and L2 acquisition are fundamentally different or not. However, this question will not be focussed in the present thesis.

42 The model accounts for differences between morphological and syntactic L2 processing as the former reveals more native like results than the latter. It additionally puts forward possible explanations of differences between child and adult processing. Since the present thesis does not intend to test the model, these implications will not be focussed any further.

Clahsen & Felser claim that L2 processing is predominantly shallow. This means that L2 learners process their L2 via the lexical-semantic route whereas deep parsing is not (fully) applied in L2 processing. Figure 2.3 above is adapted by Clahsen & Felser (2006b, p. 119) and presents their illustration of the model. Among other suggestions, the authors' propose that L2 processing might be shallow because full syntactic parsing is constrained due to the incomplete L2 grammar (again, see Figure 2.3 above).[43] Accordingly, (adult) L2 learners rely on shallower and less detailed syntactic representations than native speakers. They still are able to comprehend L2 input on the basis of lexical-semantic information since these are sufficient to chunk a sentence into meaningful units and to coordinate their semantic relationships (cf. ibid. pp.116). Consequently, this approach may account for the observation that ERPs as response to the L2 processing of semantic violations (as discussed in 2.2.1) are less different from native speakers' ERPs than those as response to syntactic violations (see 2.2.2).

Further, the model may put forward some implications about the influence of AoA and proficiency. Consider the processing of double nominative violations as described in 2.2.2: If shallow processing restricts the full activation and retrieval of syntactic processing mechanisms this would account for the different ERP patterns revealed by L1 and L2 processing. Further, if this restriction is due to incomplete L2 grammar, proficiency may show an influence in that ERPs elicited by high proficient L2 learners should be less different from L1 ERP patterns than those of low proficient L2 learners. In other words, ERPs of L2 learners as response to the processing of a syntactic anomaly should converge that of native speakers as proficiency improves. This is consistent with the idea of Steinhauer et al.'s model of temporal dynamics (see 2.2.2.4 above). However, if full parsing routines are not available to L2 processing due to

[43] Note that the authors give an extensive cross-linguistic review on the processing data of native speakers and L2 learners for numerous morphological and syntactic structures including off-line and on-line measures. They further provide a detailed description of the different demands as to the language processing system made by various linguistic structures and elaborate explanations why L2 shallow processing is sufficient for some structures but not others. Further, a great amount of their evidence does not involve results from the processing of (syntactic) violations but rather the resolution of ambiguities and long-dependencies. Since none of their reviewed data corresponds to the structures investigated in the present thesis the above interpretation of shallow processing shall be sufficient.

maturational restrictions this would indicate AoA influence, irrespective of L2 proficiency. Clahsen & Felser acknowledge potential influence of AoA and proficiency. However, they do not make any predictions on their own concerning possible proficiency influence due to improving L2 grammar and, hence, increasing automatization of according processing mechanisms (as e.g. Steinhauer et al. do). This might be due to the designs of their own studies, where L2 learners showed little variation with respect to AoA and proficiency and, further, were compared to native speakers categorically (i.e. group design). Moreover, the authors recognize that the determination of whether and to what degree a specific structure is "shallow" needs further investigations.

2.4 Summary

Chapter 2 has dealt with the differences of ERP responses between native speakers and L2 learners. Further, it was outlined that the specific linguistic structure determines the activation of neural resources and retrieval of the subsequent processing mechanisms. Studies carried out so far have shown that both AoA and proficiency of L2 learners are accountable for differences across L1 and L2 processing patterns. Furthermore, previous outcomes suggest that their influence appears rather gradual than discontinuous not only when factors were treated as continuous variables but also when they were part of a categorical design (e.g., L2 learner groups with different mean levels of proficiency or AoA). With respect to isolated semantic and syntactic L2 processing and to subsequent ERP data, previous results suggest that the semantic processing appears robustly and differences are rather small when compared with the outcomes of L1 processing patterns. Nevertheless, the relatively slight differences and explanations in the light of AoA and proficiency influence are not uniform. Syntactic processing patterns elicited by L2 learners show greater differences when compared with native speakers' ERPs. Furthermore, relative to the differences between L1 and L2 semantic processing, the reported AoA and / or proficiency influences on the observed differences reflected by ERPs as response to syntactic violations are more consistent.

The shallow structure hypothesis as proposed by Clahsen & Felser provides a good way to account for the observed differences between L1 and L2 semantic and syntactic processing. They argue that L2 processing is shallow, which is satisfactory to activate and retrieve lexical-semantic processing mechanisms but does not or to a much lesser degree suffice to process (more) complex syntactic structures. In their view, L2 processing predominantly relies on lexical-semantic cues.

The following studies do not intend to test the shallow structure hypothesis although it may be interesting to see whether e.g. proficiency has an impact on shallow processing and / or whether results provide an indication that the processing of an unlicensed NPI are different than the processing of a double nominative violation in terms of "shallow structure" determination (see above). Rather, the upcoming experiments aim to contribute more insights into the observed inconsistency of AoA and proficiency influence as a source of differences between the activation and retrieval of L1 and L2 processing mechanisms reflected by ERPs. Moreover, the following designs aim to disentangle the correlation between both factors (i.e., the older the AoA, the lower the proficiency) by including them as fixed variables into statistical analysis. Thereby, the target will be to unfold potential differences of their weighting with respect to their impact on L2 processing mechanisms. Firstly, this will be observed for isolated semantic and syntactic processing mechanisms. Hereby the ERP components N400 and P600 will be focussed, respectively. Secondly, structures that require the enhanced activation of both semantic and syntactic processing mechanisms (N400-P600 processing pattern) will be investigated in light of the formerly observed influences of AoA and proficiency. Accordingly, the present study will also address the questions whether the influences of AoA and proficiency and their weighting could also occur with the combined processing demands and whether they are similar or different. A continuity approach is assumed concerning this matter. The present design intends to trigger evidence on potentially gradual influence of AoA and proficiency. In the remainder of this thesis the methodological and procedural aspects of the conducted experiments (Chapter. 3) and their results (Chapter. 4) are presented.

3 Participants, Experimental Procedures, and Statistical Methods

Chapter 3 covers the experimental routines used for the upcoming ERP study, including participants, procedures, stimulus materials, and methods of data analysis. This ERP study comprising three experiments was conducted at the EEG Lab at the Linguistics Department of Humboldt-Universität zu Berlin. The procedure was the same for all participants.

3.1 Participants

All participants (N = 85) voluntarily took part in the study. They were recruited via PESA[44], StudiVZ,[45] and announcements on the notice boards of the language centres at Berlin and Brandenburg universities. Prior to both testing sessions (see below), the participants were asked to fill out an electronic questionnaire to evolve their handedness following the Edinburgh Handedness Design (Oldfield, 1971; samples of the

---

[44] PESA—Psychologischer Experimentalserver Adlershof ('psychological experimental server Adlershof')—is a large electronic database established by the psychological department of Humboldt-Universität zu Berlin. Any person may register who agrees to volunteer to take part in psychological and psycholinguistic experiments. One needs to provide personal information including age, status, handedness, and psychological history (pathology) (see also http://macs5.psychologie.hu-berlin.de/pesa/public/index.php [retrieved on August 20, 2012]). Invitations for experimental participation may be sent to a choice of more than 5,000 persons, where the choice is accommodated to the conditions that have to be met for the study in question. Registered persons that fit the conditions are invited electronically (by email) to participate, and they may choose to respond to invitation or not. Most (75%) native German speakers (serving as control group) were obtained through this recruitment strategy. The remaining 25% of the recruited native German speakers followed invitations on notice boards. At the time of testing only nine native Polish speakers, who met experimental preconditions, were registered in PESA—three of them followed the invitation and took part in the study.

[45] StudiVZ is a public electronic database, based on the principle of social networking, where anybody having access to the Internet may register (see http://www.studivz.net/Default, [retrieved on August 20, 2012]). StudiVZ includes a so-called group sector, where a group relates to a certain community of interest which shares the same affections, hobbies, or backgrounds, etc. One of those groups is called '*Polnische Studenten in Berlin*' ('Polish students in Berlin'). Announcements of the experiments were sent to members of this group via the social networks' messaging system. About one-third of L2 learners (n=20) were recruited through this kind of electronic means.

questionnaires are given in appendices).[46] In addition to the handedness information, the questionnaire was modified and extended to collect data on the L1 / L2 background of each participant and included further information about the following: the length of residence (LOR) in Germany, the scale-recorded frequency of German language use in accordance with different settings (viz. at university, at home, in daily life, with friends), and the scale-recorded self-evaluation of their L2 skills in reading, writing, listening, and speaking. The collection of such meta data[47] may be important for further statistical analysis. Since the current study focuses on the potential influences of L2 learners' AoA and proficiency, there is a need to control for possible further covariates as they may correlate with AoA and / or proficiency. According to LOR, it is reasonable to assume that L2 learners who have longer resided in a German speaking surrounding are being exposed to more L2 input and hence have reached higher L2 proficiency than those L2 learners who have a much shorter residence time. This may result in a strong correlation between the proficiency level and LOR. Such correlations are undesirable and therefore need to be controlled or at least to be acknowledged with respect to further analysis (see also Flege et al., 1999 for results and discussions on correlations of potential covariates with AoA). Results of the correlation tests between meta data and both AoA and proficiency are reported in Chapter 4.1.

Behavioural and EEG responses of 65 native speakers of Polish with German as their L2 (38 female) and a control group of 20 native German speakers (12 female) were recorded. According to data analysis, the participants were divided into two groups, namely native speaker and L2 learner. At the time of testing, all the participants were residents of Berlin / Brandenburg area and attended a Berlin or Brandenburg University. They were aged between 20 and 31, right-handed, and had normal vision. The experimental procedure was divided into two sessions. These two sessions did not take place on the same day. The temporal distance ranged between two days (least) and eight days (most). The first session lasted for about 45 minutes and covered the

---

46 Questionnaires for native German speakers and German L2 learners slightly differed. Native speakers did not have to report on their language acquisition histories and backgrounds.

47 The term meta data is my own usage referring to the additional linguistic background information of all participants, which was collected through the questionnaires (see text above).

introduction of the participants to the EEG-lab. Following that, the participants' level of proficiency was evaluated by recording a standard German proficiency test —the C-Test. This is a pen and paper test that consists of five cloze tests, chosen from a large variety of the standardized version of German proficiency tests provided by the language centre at Humboldt-Universität zu Berlin (see appendix 9, see also https://anmeldung.sprachen-zentrum.hu-berlin.de/cgi/ctest2.cgi?testcode=61a4b9a75bebbb3af41cc3965cdbcd1d [last retrieved on July 03, 2012]. The participants had to finish the C-Test within 25 minutes. Evaluation proceeded by the one cloze-one count principle. Proficiency was assigned following the representative levels of CEFR (see also comments in Chapter 1): 100-81 = C1-C2 (high proficiency), 80-41 = B1-B2 (low proficiency). The lowest test result was 57 (see Chapter 4.1.1 below). Along the lines of CEFR levels B1-B2 are not considered as 'low' proficient but as 'intermediate'. When the participants and the experimenters had no further questions, an appointment for the second session was scheduled. The second session comprised the EEG recording which is described in Chapter 3.2 below.[48]

3.2 Procedure of EEG Recording

The EEG recording session lasted between 90 and 120 minutes. It consisted of EOG[49] and EEG preparation, experimental run, hair wash, and short post-tests and questionnaires. Prior to EOG / EEG preparation, the participants gave signed informed consent (see appendix 10) approved by the ethics committee of the psychological department at Humboldt-Universität zu Berlin. For EEG preparation, the participants were adjusted a tight fitting cap with sewn in Ag-AgCl-electrodes (easy cap EC20, http://www.easycap.de/easycap/e/products/products.htm#15 [last retrieved August 29, 2012]). To ensure conductivity, electrodes were stimulated with an abrasive electrolyte gel (Abralyt 2000, Easycap GmbH) and filled with an electro-gel (E11

[48] In order to avoid apprehension of bias, participants were not informed about their result in the C-Test until the second session was finished.
[49] EOG = Electroocolugram

Electro-Gel, Electro-Cap International, Inc.). Impedances were reduced to < 5 kΩ. Continuous EEG was conducted in a mono-polar manner from scalp positions F7, F3, FZ, F4, F8, FC5, FCZ, FC6, T7, C3, CZ, C4, T8, CP5, CPZ, CP6, P7, P3, PZ, P4, P8, PO3, POZ, PO4, and OZ corresponding to the extended 10–20 system (Sharbrough et al., 1995, see also Figure 2.1 and comments in Chapter 2.1). Additional six electrodes were attached for reasons of signal referencing and EOG derivation. These were prepared the same way as the cap-electrodes and controlled for impedances < 5 kΩ. The signals were referenced to A1 electrode (left mastoid), and later re-referenced to the average of left (A1) and right (A2) mastoid electrodes. EOG preparation included two electrodes on the outer canthi of each eye (left and right) as well as two electrodes placed above and below the right eye, respectively. EEG and EOG were amplified by a 32-channel amplifier (Brain Vision BrainAmp DC), which was digitized at 250 Hz and recorded on a desktop computer (Dell Optiplex 740).

After EOG and EEG preparation was completed, the ongoing EEG was demonstrated to the participants. They were asked to avoid any facial or physical movements while reading and to reduce blinks to the temporal periods between sentence presentations. Next, the participants were instructed to carefully read the sentences and judge them right or wrong as per the sentences' structural correctness and semantic meaningfulness by pushing the appropriate button. The experimental stimulus material was presented on a 19" monitor (acer, AL 1923) placed on a table in 1m distance to the participant. The buttons to conduct behavioural data were positioned on the table in front of the participant.[50] Their attention was drawn to the time constraint of 3,000 ms within which their judgment was prompted following each sentence (time-out paradigm). After instruction, the participants ran a practice session of 16 sentences. During practice, the experimenter stayed next to the participant, controlling for the appropriateness of task completion. Following the practice session, the participants ran five blocks of 70 experimental sentences each. The duration of each block ranged between seven and 10

[50] Buttons were connected to the LPT-status port of the computer used to present the stimulus material. In order to avoid the handedness effects, the buttons were interchanged for every other participant, half of the participants had to push the right-hand button for 'right', and the other half had to push the left-hand button to judge 'right'.

minutes, depending on the relative speed of behavioural reactions. Between the five experimental blocks, the participants were allowed to take a break with self-determined length where they had to remain seated but were permitted to relax and / or move head and shoulders. The conductivity of electrodes was controlled during each break, ensuring impedances < 5 kΩ during the entire recording session.

After the experimental run was completed, the experimenter removed the cap with electrodes and EOG electrodes. Participants then had the opportunity to wash their hair. After that, they had to fill in a vocabulary sheet which consisted of all the words used in the experiment (see appendix 5 for a full list of stimulus items) by crossing out each word with which they were not familiar. The final task for the participants was to fill up a short post-run questionnaire asking for self-perceived length of the session and complicatedness as well as potential strategies they might have used to accomplish the task (see appendix 11). Eventually, they were paid an allowance of 25 Euros when the entire experimental run was finished.

3.3 Stimulus Material and Presentation Procedure

The whole set of experimental stimuli consisted of 3 x 45 items in accordance with the three experimental subsets—i.e., semantic incongruity, double nominative violation, and NPI-licensing (see Chapters 4.2.1, 4.3.1 and 4.4.1 for detailed descriptions and samples of the three types of sentence material, again see appendix 5 for a full list of stimulus sentences). Each item included four conditions (two acceptable and two unacceptable). All the words used (irrespective of word-class and inflection) were controlled for frequency in accordance with the electronic database 'Deutscher Wortschatz', a German corpus of words with additional frequency display.[51] Word frequency is classified into 30 subgroups and is computed in relation to the most frequent word—i.e., the lower the class, the more frequent the word. All words used in the stimulus material were

[51] See http://wortschatz.uni-leipzig.de/ for word frequency of stimulus materials. Accordingly, word frequency was last checked on December 8, 2008, when setting up the experimental program.

ranked frequency ≤ subclass 15. The overall stimulus material consisted of 540 sentences. They were pseudo-randomized and counterbalanced. Each participant saw a subset of 350 sentences (5 x 70). Randomization was set as follows: Each participant saw two or three conditions per item. It was controlled for that the conditions of the same item did not appear within the same presentation block. The stimulus sentences were programmed with the software package Presentation 12.2 (neurobehavioral systems Inc.) and presented visually. The sentence presentation was performed via a Dell Optiplex 740 computer. An asterisk '*' (presented for 1500 ms) indicated the start of a new sentence. All sentences were presented in a verbatim manner, and each word lasted on the screen for 650 ms. Noun phrases (NPs) were presented together with determiner / quantifier. The last word of each sentence was followed by a question-screen 'Richtig ? Falsch' / 'Falsch ? Richtig' ('Right ? Wrong' / 'Wrong ? Right') that remained until the participant judged the sentence, but longest for 3,000 ms. Whenever the sentence was not judged within these 3,000 ms of time constraint, a feedback screen was shown for 1,000 ms to inform the participant to try to answer faster. Following the judgement, again, the asterisk '*' indicated the beginning of the next sentence.

3.4 Statistical Methods

3.4.1 General Remarks

All final statistical analyses were carried out with the software package R 3.1.0 (R Development Core Team, 2014) on a Windows compatible PC. Two main statistical methods were used to analyse data subsets, namely Pearson correlation testing for meta data (see Chapters 3.4.2 and 4.1) and (generalized) mixed-effects modelling for behavioural and ERP data (see Chapters 4.2 to 4.4). The structures of the most complex models performed on the current datasets are listed in appendix 4. The use of mixed-effects models to statistically analyse behavioural and ERP data is a rather recent development. Therefore, some general comments on the possible advantages of these models will be issued before their exact structures regarding the present analyses are described. Furthermore, and as will be seen in Chapter 4.1 below, the factors AoA and proficiency are highly correlated, which, for model criticism, is rather problematic when multiplying them as an interaction effect.

Mixed-effects models include both fixed and random factors. Based on the principle of maximum likelihood estimates, the optimal model includes those parameters (factors) that maximize the likelihood of the sample. They have several advantages for the analysis of behavioural data, as they include the use of subject and item as crossed independent random effects, rather than nested, as assumed by traditional hierarchical models (Baayen et al., 2008). In case of analysing accuracy rates, Jäger (2008) draws on the possible advantages of generalized mixed-effects models. He infers that the outcome is the linear combination of fixed effects and conditional random effects, and is associated with subjects and items which, then, allow for the simultaneous control of the data's variance associated with subject and item in accordance with the significance of the potential fixed effects (cf. ibid. pp. 442). In accordance with this argumentation, a generalized mixed-effects model was chosen to carry out the analysis of the accuracy datasets of the current experiments (see also Chapter 3.4.3 below).

Baayen, et al. (2008) offer an insightful introduction to the use of mixed-effects models for Reaction Time (henceforth RT) data analysis. In line with Jäger (2008), they argue that such models may control for fixed effects factors and additionally for covariates bound to them simultaneously. They also state that these models do not require prior averaging of participants and items, and may cope with the behaviour adjustment of each participant. This includes the observation that each participant has his / her own speed in accordance with the experimental task—e.g., some participants may respond faster as the experimental run proceeds, while some may not be able to keep up the speed with which they started and hence decrease their response times. Thus, mixed-effects models take into account the often-observed effect of learning or fatigue by including a random intercept and a random slope for subject into the statistical model (cf. Baayen et al., 2008, p.399). A similar scenario holds for the factor item; stimulus material items (here: sentences) differ in their linguistic and contextual features, even when the stimulus material includes strictly controlled conditions. Moreover, items used in the experiment make up only a sample rather than an exhaustive list (cf. ibid. p.390). Therefore, it is rational to include item as random effect. Random slopes for item are rarely necessary and thus disregarded. One further advantage of mixed-effects models is their ability to easily deal with unbalanced data, allowing for datasets where missing data may occur (ibid. p. 396). Values of experimental conditions are weighted on the basis of their number of observations of the conditions, and the variation within and between conditions (see e.g., Gelman & Hill, 2006, pp. 529; Newman et al., 2012). Hereafter, for the analysis of the present RT data, mixed-effects models were developed (again see Chapter 3.4.3 below).

In line with the advantages claimed to analyse accuracy and RT datasets, mixed-effects models also improve the analysis of ERP data. They allow for the incorporation of both continuous and categorical predictors being fixed or time-varying and individual differences of each subject's matrix (see also related comments by Bagiella, Sloan, & Heitjan, 2000). This means that such models easily deal with non-sphericity of ERPs and hence corrections by e.g. Greenhouse-Geisser or Hyun-Feldt remain unnecessary (see Footnote 10; for related comments see also Baayen et al., 2008; and Newman et al., 2012). Thus, the mixed-effects model method was also chosen to analyse the present ERP datasets (see Chapter 3.4.4).

Irrespective of the advantage of mixed-effect models but also concerning the statistical analyses of the current behavioural and ERP data, the two following issues have to be taken into account. First, the collinearity of AoA and proficiency presents a problem. This is not easily to be solved but may slightly be facilitated by mean-centring both factors. Furthermore, leaving out one factor and estimating the effects of the other would cause a so-called omitted variable bias, which is even worse than the collinearity.[52] Therefore, inclusion of both factors into the model is supposed to reduce the bias (see e.g., Clarke, 2005; Resmeth, 2012, November 28, my variables have an

[52] With respect to the omitted variable bias, there is the central assumption that the effect of a correlation of two variables is not separable to either one variable. However, a concrete theoretical interpretation is not available yet (F. Golcher, personal communication Feburary 11, 2015; see also Wurm & Fisicaro, 2014).

unhealthy relationship [Web log comment], retrieved from http://blogs.lse.ac.uk/methodology/2012/11/28/unhealthy/). Second, native speakers do not include any variance relative to the AoA factor. All reported AoA = 0 (birth). Therefore, the statistical models performed on the native speaker's data excluded AoA as a factor. Additionally, as will be seen below, the factor proficiency also shows only very little variance in native speakers. Given the relatively small group (compared with L2 learners) and the small variance in proficiency, the latter also did not enter statistical models performed on native speakers' data.[53]

The following subchapters provide detailed information on the used statistical functions. The level of significance is consistently set at $p < .05$.

3.4.2 Meta Data

Given the sample variances due to collected meta data for all relevant averages (means of subset data by all participants), single *correlation tests* with variables (factors) AoA and proficiency were applied following *Pearson's Product-Moment Correlation Coefficient*[54]. In order to find out whether AoA or proficiency better relate to any of the meta data's variables, the two relevant overlapping correlations based on dependent groups were compared by performing function *concor()*, R-package 'cocor' (Diedenhofen, 2013).

---

[53] Pakulak and Neville (2010) report changes in ERP responses to English insertion violations (e.g., stronger P600) as a function of higher L1 proficiency. The proficiency of 34 native speakers was related to the average standardized score for three subtests of the TOAL-3. These averages significantly differed between high and low proficient groups (n = 17, respectively). Additionally, the participants' working memory performance was controlled. Also, Newman et al. include the proficiency score of native speakers as a within-group factor. Their data shows much more variation in the mean proficiency scores between native speakers. In the present analysis of native speakers' data, however, proficiency has not been included due to little variation and relatively small number of participants.

[54] Pearson's Product-Moment Correlation Coefficient measures the strength of the linear dependency of two variables. It is denoted by r and ranges from -1 to 1, where a 0-value indicates no correlation, 1 the perfect positive, and -1 the perfect negative correlation. Although the data of both AoA and proficiency were not normally distributed, the Pearson's method was chosen. Additionally, a test of Spearman's Rho was performed on all datasets by using the function *rcorr()*, R package 'Hmisc' (Harrell, 2014). Yet, differences pertaining to the results of the Pearson's Product-Moment Correlation Coefficient are negligible and therefore not reported.

3.4.3 Behavioural Data: Accuracy Rates and Reaction Times

The main difference between accuracy rates and RTs is that the outcomes of the former are categorical in nature, whereas those of the latter are continuous. Analyses of categorical data require the use of a generalized linear mixed-effects model, whereas RTs are applicable to linear mixed-effects models. Both methods include the search for an optimal model that best describes the relevant data.

Analyses of accuracy data were carried out with function *glmer()* (R-package 'lme4', Bates, Maechler, Bolker, & Walker, 2014). Accuracy ratings served the model as dependent variable. Initially, the most complex model included condition[55] (within-subject), AoA, proficiency (both between-subject), and all possible interactions as fixed effect factors. The random structure integrated by-subject and by-item random intercepts as well as by-subject random slopes for condition. Random slopes for AoA and proficiency were not included as they are not within-subject factors (see appendix 4.1 for the full (most complex) model). To find the optimal model, a stepwise reduction of parameters was performed first on random-effects and then on fixed-effects. The optimal model then was chosen through the minimum Akaike Information Criterion (AIC, see Akaike, 1973, 1974; see also Burnham & Anderson, 2004).

Prior to the analyses of present RT data, only physically impossible short RTs (<170 ms) were removed in accordance with the method suggested by Ratcliff (1993). Because of the timeout long latencies did not exceed 3,000 ms. Possible long outliers were not removed from original RTs.[56] Density estimates for all conditions were not normally distributed. Therefore, following Baayen and Milin (2010), RTs were log-transformed, allowing that log-normal distributions of all conditions increased goodness over original distributions. In the beginning, the most complex model was determined. This included the logarithmized reaction times (RT.log) as dependent variable as well as condition (see

[55] Condition is used here as the umbrella term and refers to the specific linguistic contrasts—i.e., Congruity (in Chapter 4.2), Case (in Chapter 4.3), and Licensor (in Chapter 4.4), respectively. These latter terms are capitalised whenever they refer to the factor within the statistical model.

[56] According to Ratcliff (1993), it is hard to locate and isolate long outliers. Although they might increase means, they may remain in the model and be treated by model criticism, rather than isolated by a-priori screening methods (see also Baayen & Milin, 2010, pp. 19).

Footnote 55), AoA, proficiency, and all possible interactions as fixed effect factors. By-subject and by-item random intercepts and by-subject random slopes for condition were included as random effects. The structures of the full / most complex models are listed in the appendix 4.2. To find the best fitting model, the following bipartite procedure was carried out. In the first step the optimal model was determined. It means that progressively simpler structured models iteratively were compared with more complex ones using log.likelihood ratio testing (function *lmer(),* R-package 'lme4', Bates et al., 2014). Next, by using R package 'LMERConvienience Functions' (Tremblay & Ransijn, 2013), outliers were removed from the optimal model (function *trim.data.frame()*), which then was refitted (function *fitLMER()*) for both random (forward-fitting) and fixed effects (back-fitting). Residuals of the refitted models reached normal distribution. For estimation of the denominator degrees of freedom (dfs in the following), both upper and lower bound values were calculated.[57] Those values were similar since the numbers of data points of each case were relatively large. In case of significant terms of higher order, these were plotted for visual purposes by using function *Effect()*, R package 'effects' (Fox, 2003; Fox & Hong, 2009).

3.4.4 ERP Data

The software package Brain Vision Analyzer (Brain Products) was used for pre-processing of EEG data. ERPs were time-locked to the critical item of each condition (segments) and had a duration of 1,500 ms. All segments were calculated to a 100 ms pre-stream baseline yielding a total length of 1,600 ms. ERPs were filtered (low cutoff 0.05 Hz) and corrected for artefacts (e.g., blinks, muscle and facial movements as well as alpha waves). Due to high amounts of artefacts within EEG signals (native speakers [1], L2 learners [5]) and technical problems during recording (native speakers [1]), the

---

[57] As explained by Newman et al. (2012, p. 1,209), '*df* to compute upper bound probability values were calculated as the number of data points minus the number of *df* used up by the fixed effects. Those for lower bound *p* values […] minus the number of *dfs* used up by fixed effects and the number of random effects […].'

data of these seven participants had to be excluded from further analysis (see also below). For the remaining data (13,719 trials), averages were calculated for individual electrodes, including all segments per condition, for each participant. Grand averages for each electrode were calculated for the averages (of each condition) for native speakers and three groups of L2 learners.[58] For visual purposes, a 15 Hz high cut-off filter was processed on the grand average data. To enhance the readability of ERP plots, electrodes were grouped into six regions of interest (ROIs).[59] Data points (i.e., average potentials of all electrodes for each condition and each participant) in the appropriate time-windows (see Chapters 4.2.3, 4.3.3, and 4.4.3) were exported for further statistical analysis. Similar to the procedure carried out on RT data (see Chapter 3.4.3 above), statistical analyses of the present ERP datasets were performed in two steps. First, the most complex model was determined (function *lmer()*).[60] This model included the mean potential as dependent variable and the following fixed effects factors: condition (again see Footnote 55), ROI, AoA, and proficiency and all possible three and two-way interactions. The random structure included by-subject intercepts as well as by-subject random slopes for condition and ROI (see also appendix 4.3 for the most complex structure of the corresponding statistical model). The second step (using package LMERConvenienceFunctions, for references see Chapter 3.4.3 above) comprised removal of outliers (function *trim.data.frame()*) and the refitting of the model for both random and fixed effects (function *fitLMER()*). Residuals of all cases of refitted models were normally distributed. Denominator *dfs* of both upper and lower bound values were calculated (again see Footnote 57). Finally, if the terms of higher order of the fixed effects were reliable, the function *Effect()* was performed to plot these higher-ordered

---

58 For visual purposes only, grand averages of L2 learners were grouped on the basis of their AoA and proficiency level (see also Chapters 4.2.3, 4.3.3, and 4.4.3). The model used for statistical analyses of L2 data did not include group as a factor.

59 ROIs for visual purposes were arranged by a 2x3 design (anterior / posterior x left / mid / right): left-anterior = F3, F7, FC5, C3; midline-anterior = Fz, FCz, Cz; right-anterior = F4, F8, FC6; C4; left-posterior = P3, PO3, P7, CP5; midline-posterior = CPz, Pz, POz; and right-posterior = P4, PO4, P8, CP6.

60 The structure of the most complex statistical model used here is adapted from Newman et al., 2012. It had to be slightly modified in accordance with the fixed effects structure (including AoA as a continuous fixed effect) and was approved by Antoine Tremblay (A. Tremblay, personal communication, November 22. 2012).

terms, respectively. Significant interactions, including ROI, were further resolved by *post hoc* analysis of single ROIs. Additionally, and only if necessary, to gain more insights into potential differences between the two conditions of a contrast (see Footnote 55), significant interactions were also resolved for the factor condition.

4 Results

This chapter reports the behavioural and ERP results of the three experiments conducted to investigate L1 and L2 processing of semantic (in-) congruity (4.2), double nominative violation (4.3), and NPI licensing conditions (4.4). Before the results are presented, the statistical correlations between AoA and proficiency as well as their statistical relation to the subsets of meta data will be presented in detail in Chapter 4.1 below.

4.1 Meta Data

Meta data includes all the background information which has been collected from the participants (N = 78)—i.e., native speakers (n = 18) and L2 learners (n = 60)—via the electronic questionnaires (see Chapter 3.1; samples of questionnaires are given in appendices 7 and 8). Generally, all L2 learners report German as their L2 and additional knowledge of a third language (L3)—either English or Russian or French. L2 learners (n = 11) who acquired German and Polish rather simultaneously (between age 0 and 4 according to Meisel, 2011)[61] report that Polish has always been their dominant language as was the case for the remaining L2 learners.

With respect to the meta data analysis AoA and proficiency (C-Test scores) were statistically correlated to find out whether both factors enter into a significant relationship. Also, the data on self-evaluation of linguistic L2 skills, amount of L2 usage, L2 vocabulary knowledge, and LOR, taken from the questionnaire (again, see appendices 7 and 8), were statistically correlated with AoA and proficiency to measure the strength of both factor's prediction on these separate meta datasets, respectively. Correlation tests were performed globally—i.e., including all data points (L1 and L2

[61] Again, the AoA distinction between simultaneous and successive L2 acquisition should not be attributed to me. It is adapted from Meisel, 2011.

data)—and also only for L2 data. In the following, results are reported separately (for the description of statistical methods, see Chapter 3.4.2.).

4.1.1 Proficiency and Age of Acquisition

The overall proficiency elicited by C-Test scores is high (89.14%). On average, native speakers reach 96.6% (range: 93% to 99%) and L2 learners reach 87.19% (range: 57% to 99%). Figure 4.1 illustrates the rank-ordered scores of the C-Test for each participant. Figure 4.2 displays the individual AoA for each participant. All native German speakers report birth as their AoA. L2 learners' AoA ranges from 0 to 20 years (mean = 9.37 years).

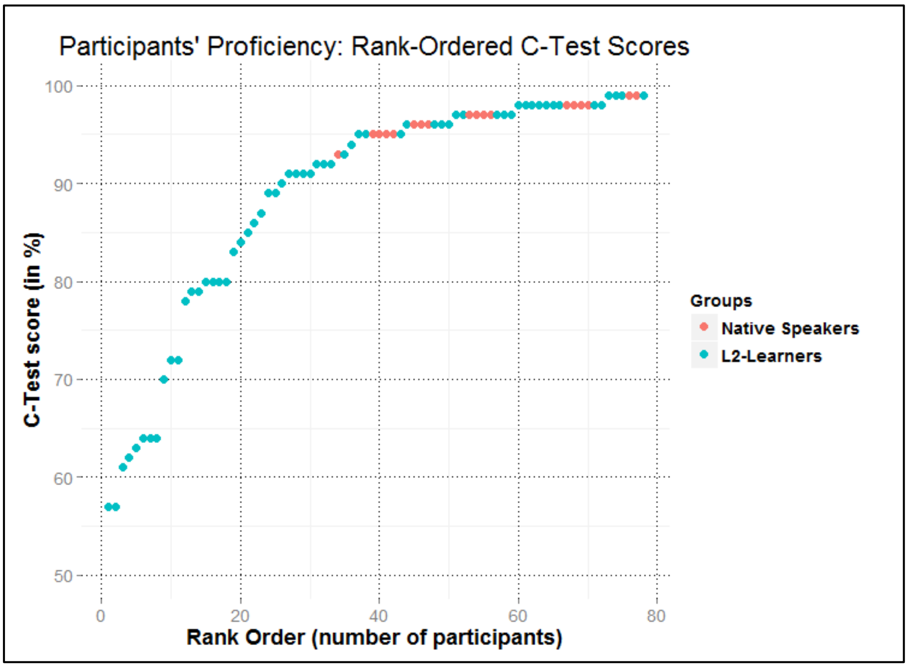


Figure 4.1: List of participants (n = 78) are mapped on x-axis rank-ordered by C-Test scores in % (y-axis). Native speakers = orange, L2 learners = blue.

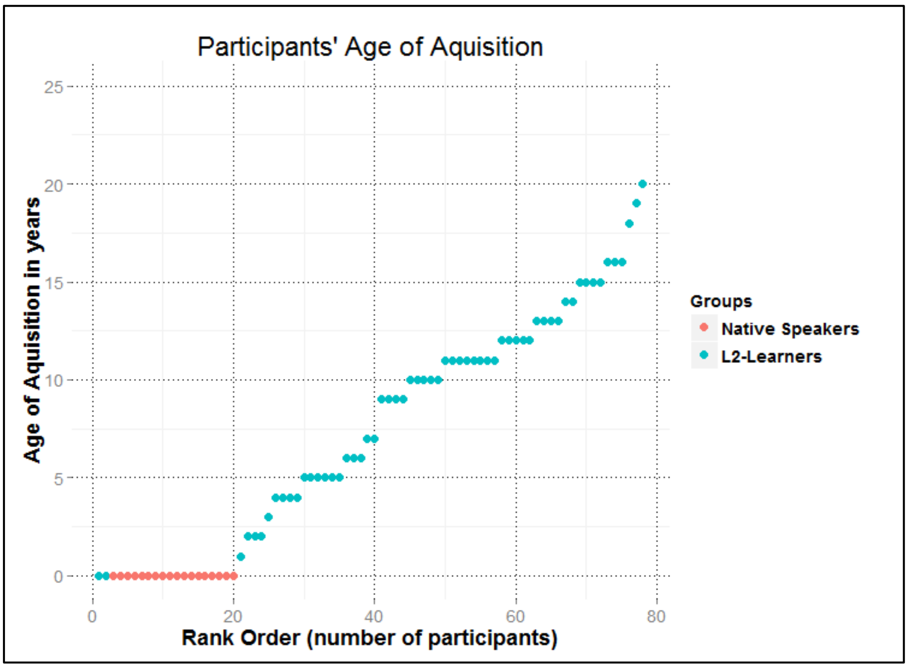


Figure 4.2: List of participants (n = 78) mapped on x-axis ordered by AoA in years (y-axis). Native speakers = orange, L2 learners = blue.

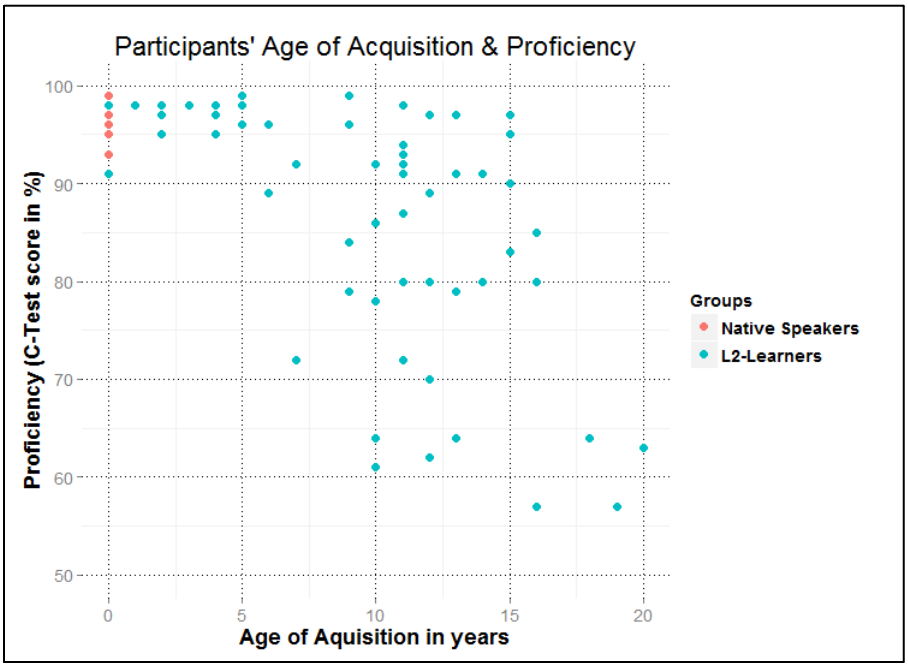


Figure 4.3: AoA in years (x-axis) and proficiency according to C-Test score in % (y-axis) illustrated for all participants (n = 78). Native speakers = orange; L2 learners = blue. Identical values (e.g., AoA = 5, proficiency = 99) are mapped only once.

Results of the global *post hoc* test (Pearson) to find out about the correlation between the proficiency and AoA of all participants reveals a highly significant negative correlation $r = -0.64$, which is illustrated in Figure 4.3. The correlation indicates that as AoA increases, proficiency decreases that much ($t(76) = -7.34$, $p < .001$). A further correlation test was performed by taking data from only L2 learners. Similarly, the outcome reveals a highly significant negative correlation between AoA and proficiency ($r = -.59$, $t(58) = -5.63$, $p < .001$).

4.1.2 Self-Evaluation of German Language Skills

Figure 4.4 illustrates the average of the participants' self-report of L2 German language knowledge on the basis of their self-evaluation in linguistic skills such as reading, writing, listening, and speaking. Self-evaluation was conducted by using a scale ranging from '0' to '5', with '0' indicating 'no knowledge' and '5' associated with 'native'. The questionnaire for native speakers did not include the self-evaluation of skills in their L1 (German). All of them (n = 18) were given the highest rate ('5 = native knowledge') for each skill. Thus, the mean value with which their data was entered into statistical analysis is '5'.

Post hoc tests (Pearson) were performed to test the correlation between the average of the self-evaluation and AoA and proficiency. Results show that both correlations are highly significant (~proficiency ($t(76) = 13.7$, $p < .001$, $r = .84$); ~AoA ($t(76) = -11.9$, $p < .001$ $r = -.8$). This indicates that individual scale ratings for the self-evaluated skills in L2 German are rated higher as proficiency improves and deteriorates with increasing AoA. Additional tests for only L2 data indicate similar correlations with proficiency ($t(58) = 11.53$, $p < .001$, $r = .83$) and AoA ($t(58) = -7.66$, $p < .001$, $r = -.7$). *Post hoc* comparisons testing the statistical difference between the respective correlations reveal that proficiency better relates to the self-reported L2 German skills than AoA (Pearson: $z = 28.26$, $p < .001$ for all data points; Pearson: $z = 18.26$, $p < .001$ for only L2 learner data points). The high correlation between the self-reported evaluation and proficiency indicates that L2 learners' self-perceived proficiency largely matches C-Test scores.

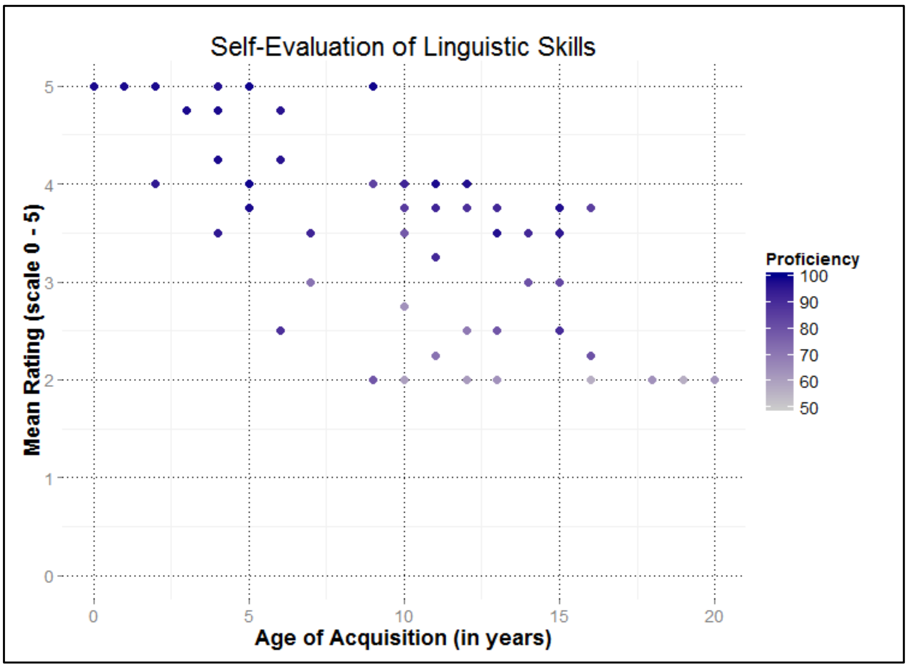


Figure 4.4: Average of participants' (n = 78) reported self-evaluation of German linguistic skills in reading, writing, listening and speaking on a scale between 0-5 (scale ratings indicate 5 = native, 4 = fluently, 3 = good, 2 = intermediate, 1 = basic, 0 = no knowledge). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, mean rating = 5 and shading = 90-100) are mapped only once. Mean rating of all native speakers (n = 18) = 5 (with AoA = 0).

4.1.3 Self-Reported German Use

Figure 4.5 illustrates the average of L2 learners' self-estimation on how much / often they are using their L2 German in various situations of daily life—i.e., 'daily life', 'at work / university', 'with friends', and 'with family members'. At the time of testing, all the participants had lived in Berlin or Brandenburg area and mastered their daily lives in a predominantly German-speaking surrounding. The present outcomes of *post hoc* correlation tests (Pearson) reveal a significantly positive correlation between the self-reported German use and proficiency ($t(76) = 7.38$, $p < .001$, $r = .64$) and a negative correlation with AoA ($t(76) = -9.11$, $p < .001$, $r = -.72$). This indicates that the amount of self-reported German use increases as proficiency level improves. AoA influence is obvious through decreasing L2 usage. Similar outcomes are revealed by correlation tests with the data derived from only L2 learners—i.e., ~proficiency ($t(58) = 5.69$, $p < .001$, r

= .59) and ~AoA (*t*(58) = -4.14, *p* < .001, *r* = -.47). The *post hoc* comparisons of the two types of correlations reveal that proficiency better relates to the amount of self-reported German use than AoA (Pearson: *z* = 13.43, *p* < .001for all data points; Pearson: *z* = 6.84, *p* < .001 for only L2 learner data points).

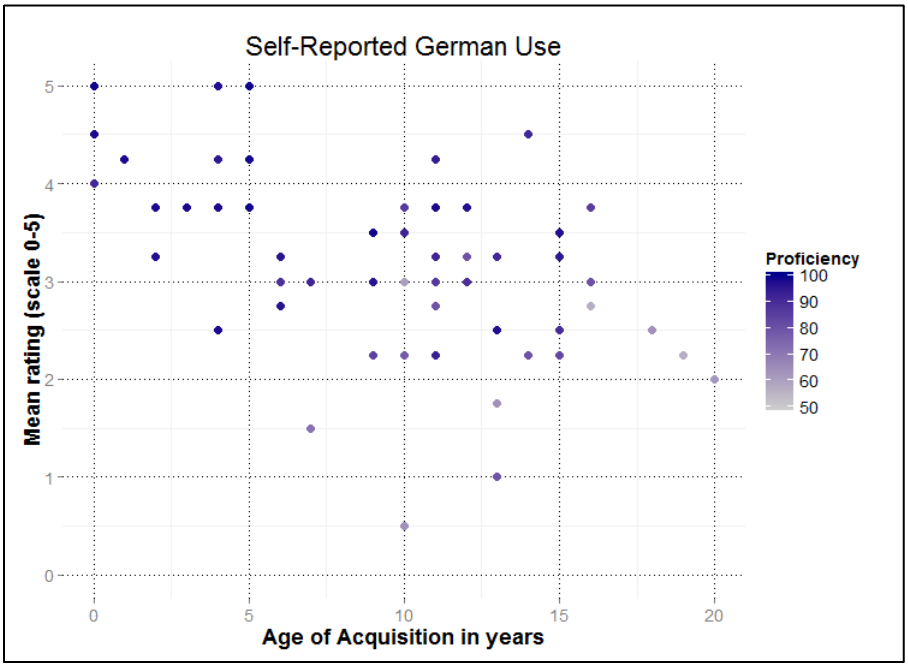


Figure 4.5: Average of participants' (n = 78) self-reported German use in four different communicational situations as "daily life", "at work/university", "with friends" and "with family members" (scale ratings indicate 5 = always, 4 = most of the time, 3 = sometimes, 2 = seldom, 1 = only when necessary, 0 = never). X-axis displays AoA. Shading of points indicates individual proficiency level i.e. the measured C-Test score. Identical values (e.g., AoA = 0, mean rating = 5 and shading = 90-100) are mapped only once. Mean rating of native speakers (n = 18) = 5 (with AoA = 0).

4.1.4 Self-Reported Vocabulary Knowledge

Figure 4.6 illustrates the vocabulary knowledge reported by each participant. The overall knowledge of vocabulary used for stimulus material is high. The mean vocabulary knowledge of all participants (n = 78) is 93.51%, for only L2 learners (n = 60) it is 91.56% (range 57 to 100). *Post hoc* correlation tests (Pearson) yield a significant positive correlation between vocabulary knowledge and proficiency (*t*(76) =11.88, *p* < .001, *r* = .8), indicating an increase in vocabulary knowledge as proficiency is improved, and a significant negative correlation between vocabulary knowledge and AoA ($t(76) = -7.53$, $p < .001$, $r = -.65$) that indicates a higher vocabulary knowledge as AoA is early. The calculation of only L2 data largely resembles the values of the global correlations with proficiency ($t(58) = 9.45$, $p < .001$, $r = .77$) and AoA ($t(58) = -5.43$, $p < .001$, $r = -.58$). The comparisons between the respective correlations, again, show that proficiency better predicts vocabulary knowledge than AoA (Pearson: $z = 16.5$, $p < .001$ including all data points; Pearson: $z = 11.82$, $p < .001$ for only L2 learner data points).

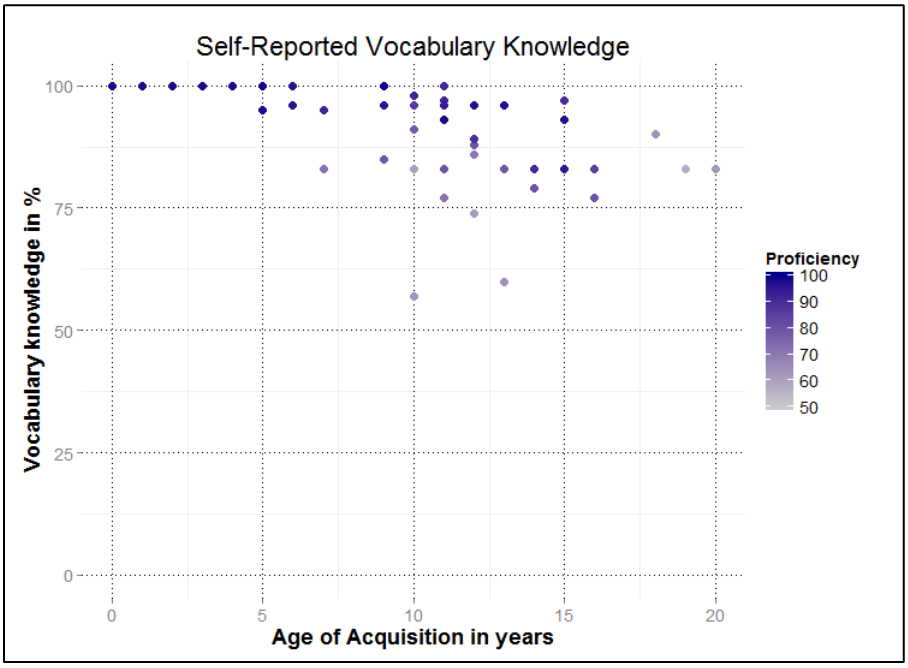


Figure 4.6: Average of self-reported vocabulary knowledge obtained for all participants (n = 78) according to the words used in the stimulus material. X-axis displays AoA. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, vocabulary knowledge = 100% and shading = 90-100) are mapped only once. Mean vocabulary knowledge of native speakers (n = 18) is 100 % (with AoA = 0).

4.1.5 Length of Residence

As already indicated, at the time of testing, all the participants were residents of Berlin or Brandenburg Area and attended some local institution of higher education or had already gained an academic degree. Therefore, it may be assumed that all participants have comparable educational qualifications.

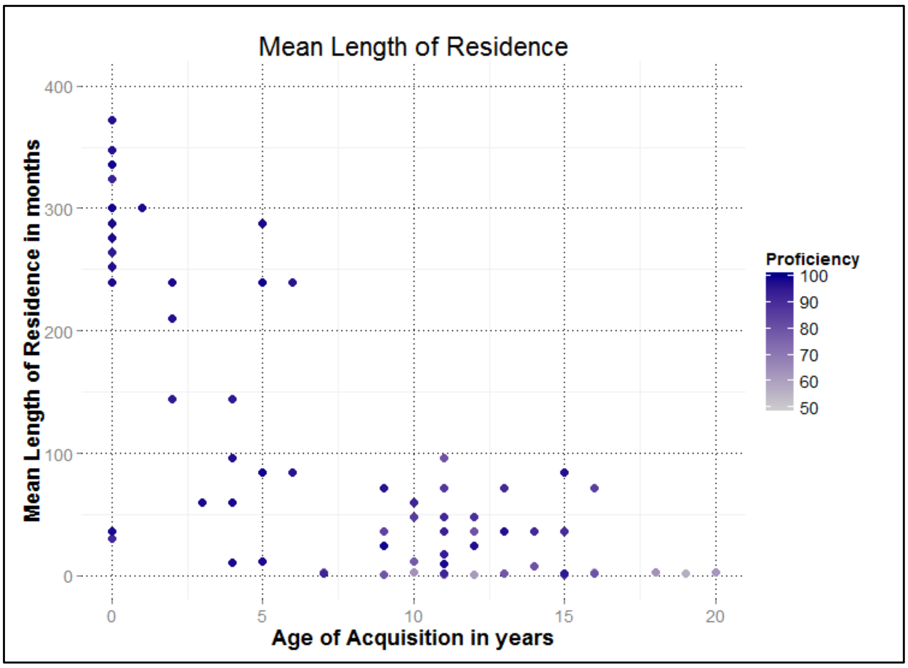


Figure 4.7: Participants' (n = 78) mean length of residence (LOR) in months based on their time residing in Germany. X-axis displays AoA. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, LOR = 300 and shading = 90-100) are mapped only once.

The mean periods during which the participants reported residence in a German-speaking surrounding are depicted in Figure 4.7. Note that, for native German speakers, the exact age (given in months) at the time they were tested is considered as their mean LOR (294.06). *Post hoc* correlation tests (Pearson) revealed a high positive correlation between LOR and proficiency ($t(76)$ =5.69, p < .001, r = .54), and a strong negative one between LOR and AoA ($t(76)$ = -11.01, p < .001, r = -.78). The outcomes are similar when only L2 learners' data is considered (~proficiency ($t(58)$ = 4.22, p < .001, r = .48) and ~AoA ($t(58)$ = -5.23, p < .001, r = -.56)). *Post hoc* comparisons of the correlations reveal significant differences that show AoA rather than proficiency to relate to LOR (Pearson: z = 12.44, p < .001, includes all data points; Pearson: z = 6.54, p < .001 includes only L2 learner data points).

4.1.6 Summary: Meta Data

The previous correlations tests have shown two things: (i) AoA and proficiency are highly correlated and (ii) *post hoc* correlation tests of both factors with subsets of meta data reveal highly positive correlations with proficiency and highly negative ones with AoA, irrespective of whether all data points or only L2 learners' data entered the statistical analyses. Given this similar influence of both AoA and proficiency on all the covariates, they will not be issued as independent factors. Furthermore, and with respect to German skills, use, and vocabulary knowledge, L2 learners' ratings seem to be better predicted by proficiency than by AoA. This is considered desirable and supports the idea to treat proficiency as an independent factor.

The following three subchapters provide the results of behavioural and ERP data as response to L1 and L2 processing. As already indicated, both AoA and proficiency are treated as independent variables. It is their very influence on L2 processing phenomena that make up the primary subject of investigation. As was shown above, both these factors are highly correlated—increasing AoA yields lower proficiency. Usually, when two factors show a significant correlation, one factor has to be controlled to reliably reveal the potential influence of the other on any further dependent variable. In the following models, neither factor will be controlled, as both will enter the models as independent and continuous variables with mean-centred values. The above-mentioned comparisons between the correlations reveal that there are differences in the predictive strength of both factors. Therefore, for the following analyses, it will be of main interest whether the processing results elicited by ERPs, in particular, also reveal differences in the predictive strength between AoA and proficiency. In other words, it will be investigated which factor better relates to the potential differences between L1 and L2 ERP patterns and whether these relations vary concerning the structure that is processed.

4.2 Experiment 1: Semantic Incongruity

4.2.1 Stimulus Material and Hypotheses

This first experiment investigates the potential differences between L1 and L2 processing of semantic incongruity and whether AoA and / or proficiency have an influence. Semantic incongruity is restricted to the sentence level context. Sample stimulus sentences are displayed in Table 4-1 (see appendix 5.1 for a full list of stimulus sentences). All incongruent sentences contain a strong semantic violation with respect to the sentence context (rather than a violation of world knowledge, see Chapter 2.2.1.1).[62]

Table 4-1: Stimulus examples of semantic congruent and semantic incongruent sentences: NOM = Nominative, ACC = Accusative, '*'-indexed sentences are incongruent

Correct congruent condition Der$_{\mathbf{NOM}}$ Autor$_{\mathbf{NOM}}$ schreibt *den*$_{\mathbf{ACC}}$ *Roman*$_{\mathbf{ACC}}$ und erhält einen Preis. *The*$_{\mathbf{NOM}}$ *author*$_{\mathbf{NOM}}$ *writes the*$_{\mathbf{ACC}}$ *novel*$_{\mathbf{ACC}}$ *and wins a prize.*
Incorrect incongruent condition *Der$_{\mathbf{NOM}}$ Autor$_{\mathbf{NOM}}$ schreibt *den*$_{\mathbf{ACC}}$ *Stuhl*$_{\mathbf{ACC}}$ und erhält einen Preis. **The*$_{\mathbf{NOM}}$ *author*$_{\mathbf{NOM}}$ *writes the*$_{\mathbf{ACC}}$ *chair*$_{\mathbf{ACC}}$ *and wins a prize.*

The linear word order of this set of stimulus material is 'S V O *and* V O'. This word order represents the unmarked / canonical word order of the matrix sentence structure and resembles both German and Polish (but see Footnote 23). The subject of each sentence is a nominative, singular NP, which is further marked as masculine, definite, and animate. The subject precedes a transitive verb, and the subject-verb agreement is realized, given all appropriate features. The verb is followed by the critical item, which

[62] Additionally, two conditions were presented. These two additional conditions included a subject quantifier NP 'Kein Autor' ('no author'), instead of a subject determiner NP 'Der Autor' ('the auhor'). Inclusion of these two additional conditions was necessary because of the structure of the stimulus material used for the third experiment on NPI licensing (see Chapter 4.4.1) and thus to prevent the participants from building up predictive strategies based on the similarity of the structure of stimulus sentences. For analysis, both additional conditions are treated as fillers and are not observed any further.

is an object NP-marked accusative, singular, masculine, as well as definite and non-animate. Semantically, the critical item either matches the prior context (congruent condition) or does not (incongruent condition). All sentences continue with the conjunction 'und' ('and') followed by a second transitive verb and a direct object NP also marked as accusative, singular, masculine, non-animate but indefinite. The continuation serves to prevent sentence wrap-up effects.[63] All stimulus sentences are syntactically well formed. Acceptability rejections of half of these sentences should be due to semantic violation expressed on the critical item. As soon as the language processor encounters a direct object that does not fit the prior context semantically, the processing costs are enhanced due to integration problems. These integration problems are revealed by the mismatch with former prediction. This mismatch hinders the updating processes, which should be reflected by an enhanced N400 (again, see Chapter 2.2.1.1).

With respect to prior ERP studies on the processing of semantic incongruity, an enhanced N400 component as response to semantic incongruent compared with semantic congruent words in sentences—i.e., the N400 effect—is expected. Furthermore, this N400 effect should be elicited by the ERPs of both native speakers and L2 learners. There might also be differences across L1 and L2 ERP patterns concerning the strength, distribution, and latency of the N400 effect. Whether these differences will be due to AoA and / or proficiency influence cannot be clearly hypothesized on the basis of previous findings (again, see Chapter 2.2.1.2). The influence of decreasing proficiency might be reflected by gradual attenuation and / or increasing left-lateralisation of the N400 effect (Ardal et al., 1990; Newman et al., 2012; Ojima et al., 2005). However, attenuated strength as well as lesser right-hemispheric dominance of the N400 effect might also be caused by AoA influence (Hahne, 2001; Weber-Fox & Neville, 1996).

[63] Sentence wrap-up effects may occur at the end of a sentence comprising processing mechanisms that reflect integrative processes concerning the entire structure (see e.g., Just & Carpenter, 1980 for general comments; for short comments on wrap-up effects in ERP research see also Rispens & Krijkhaar, 2010). It may therefore be the case that processing effects revealed by critical items that occurred at the end of a sentence may be spurious to interpret since they entail higher processing loads which might be impossible to isolate and hence be referred exclusively to be enforced by the critical item.

Recall from Chapter 2.2.1.3 that the questions to be investigated refer to (i) whether and (ii) how AoA and proficiency may impact the L2 processing of semantic incongruity and hence the N400 effect. With respect to the statistical design of the present study, both AoA and proficiency will be directly linked to the strength, distribution, and latency of the expected N400 effect. Further, to see whether the influence of both factors appears rather gradual than discontinuous, as is expected, they are treated as continuous variables.

4.2.2 Behavioural Results: Semantic Incongruity

Mean accuracy rates obtained for all participants are presented in Figure 4.7 and Table 4-2. Statistical analyses of native speakers' accuracy judgments do not yield significant differences between congruent or incongruent sentences (model only had an intercept). L2 learners' accuracy judgements for the congruent sentences are slightly higher than for the incongruent sentences, which, however, is statistically not reliable ($F(1,3384) = 2.73$, $p = .09$)[64]. Both improving proficiency and increasing AoA influence the rating of L2 learners' mean accuracy. The former significantly enhances accurate ratings ($F(1,3384) = 74.2$, $p < .001$)[65]. Increasing AoA reliably decreases the accuracy of judgements ($F(1,3384) = 17.27$, $p < .001$)[66]. Both factors also significantly interact ($F(1,3384) = 6.29$, $p = .01$),[67] but this interaction does not affect Congruity. At early AoA, improving proficiency largely enhances accuracy; this influence gradually reduces as AoA increases. At latest AoA stages, the level of L2 proficiency does not yield any influence.

---

[64] Throughout the thesis results are accordingly reported for *denominator upper bound*. Footnotes index the according *denominator lower bound* = 3174, $p = .09$.
[65] *Denominator lower bound* = 3174, $p < .001$.
[66] *Denominator lower bound* = 3174, $p < .001$.
[67] *Denominator lower bound* = 3174, $p = .01$.

Table 4-2: Mean accuracy rates (in %) and RTs (in ms) revealed by native speakers and L2 learners for semantic congruent and incongruent sentences with standard deviations in parentheses.

Group	Accuracy in %		Reaction Times in ms	
	congruent	incongruent	congruent	incongruent
Native Speakers (n = 18)	92.78 (25.89)	89.03 (31.27)	496.45 (300.49)	498.92 (269.01)
L2 learners (n = 60)	73.38 (42.48)	68.16 (46.59)	705.28 (472.27)	671.35 (438.91)

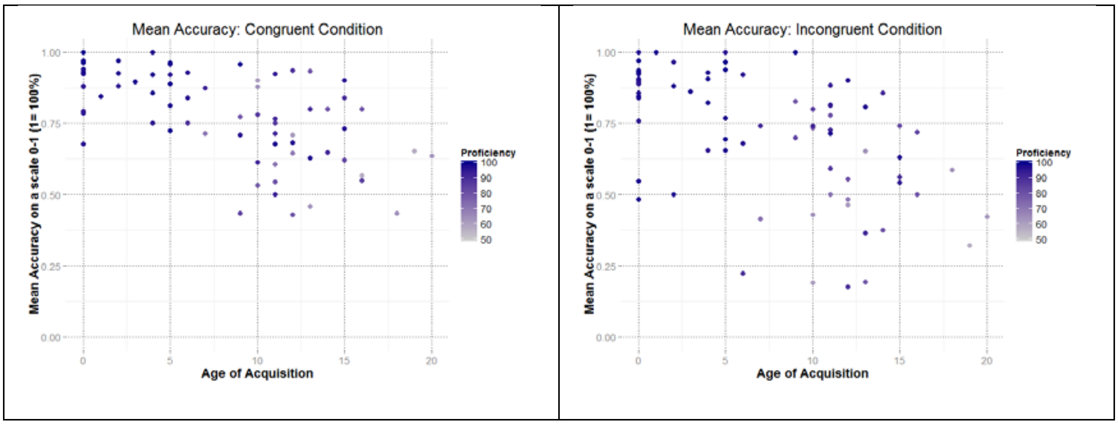


Figure 4.8: Mean accuracy rates (in %) for congruent (left) and incongruent (right) conditions for all participants (n = 78). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, accuracy = 100% and shading = 90-100) are mapped only once.

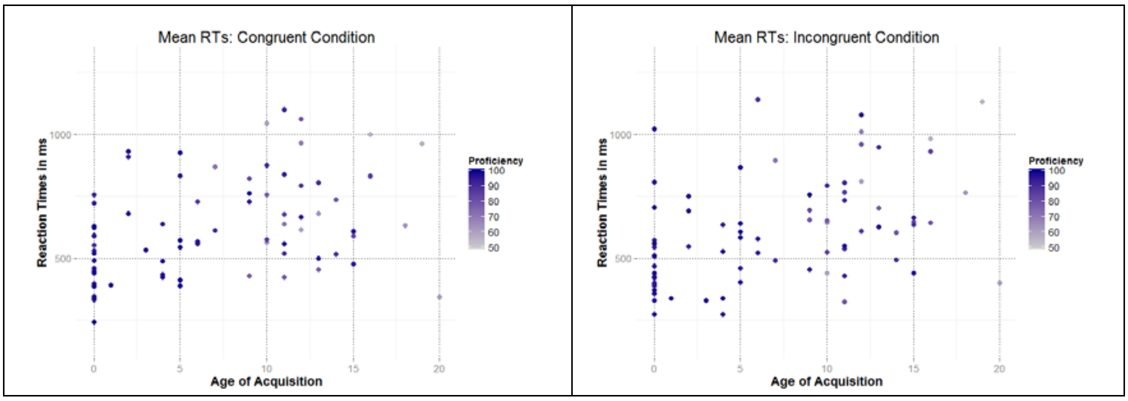


Figure 4.9: Mean reaction times (in ms) for congruent (left) and incongruent (right) conditions for all participants (n = 78). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, RT = 500 ms and shading between 90-100) are mapped only once.

Average RTs of the accurately judged congruent and incongruent sentences are displayed in Figure 4.8 (above) and also given in Table 4-1 (above).[68] Native speakers do not show any significant speed differences between the ratings of congruent and incongruent sentences (model only had an intercept). The statistical results for L2 learners do not reveal any reliable main effects for Congruity ($F(1,2379) = 1.01$, $p = .3$)[69] and AoA ($F(1,2379) < 1$). Yet, due to increasing AoA, RTs differ in terms of Congruity ($F(1,2379) = 4.56$, $p = .05$)[70]. Increasing AoA gradually slows down the RTs of incongruent sentences, whereas it fastens the speed to accurately judge a congruent sentence. Improving proficiency reliably speeds up the judgments of both congruity conditions ($F(1,2379) = 3.87$, $p = .05$)[71]. Since the critical item is far removed from the end of the sentence and ratings might also be influenced by sentences' continuation, accuracy and RT data are seen as offline measures and will not be issued in further interpretations.

4.2.3 ERPs: Semantic Incongruity

Overall, ERPs of the semantic conditions included 4,564 trials for averaging. Due to blink and drift-artefacts 695 trials (15.22%) had to be excluded from analysis. Figure 4.9 illustrates the grand average difference ERP waves for native speakers and L2 learners.[72] ERP waves are depicted in a 1,600 ms time-window. The presentation of the critical item is linked to 0 ms. Electrodes are pooled into six ROIs in accordance with their location on the scalp (see Chapter 3.4.3 and Footnote 59). Voltage difference maps are presented by Figure 4.10. They illustrate the mean voltage difference between the ERP

---

[68] According to the accuracy rates elicited by L2 learners, a data loss of approx. 27% has to be considered when removing all inaccurately rated sentences. Therefore, RT data was additionally analysed, including all data points in the model: This additional analysis only reveals a significant main effect for proficiency ($F(1,3381) = 3.93$, $p = .04$, reported for denominators upper bound), indicating faster judgments as proficiency improves. AoA does not show any influence.
[69] *Denominator lower bound* = 2187, $p = .3$.
[70] *Denominator lower bound* = 2187, $p = .05$.
[71] *Denominator lower bound* = 2187, $p = .05$.
[72] Grand average difference was calculated on the basis of the mean voltage of the congruent condition subtracted from the mean voltage of the incongruent condition.

waves of the two conditions in a 400–500 ms time-window. For sole visual purposes L2 learners are grouped on the basis of their AoA—i.e., early (< 8) vs. late (≥ 8)[73]—and their proficiency level which was determined by the C-Test score (high (≤ 80%) vs. low (> 80%))[74]. This yields three L2 learner groups, namely: early L2 learners with high proficiency (EAHP), late L2 learners with high proficiency (LAHP), and late L2 learners with low proficiency (LALP). Importantly, the linear mixed effects model does not include *group* as a factor. Single group ERPs illustrating the comparison between both congruent and incongruent conditions are given in appendix 1.1.

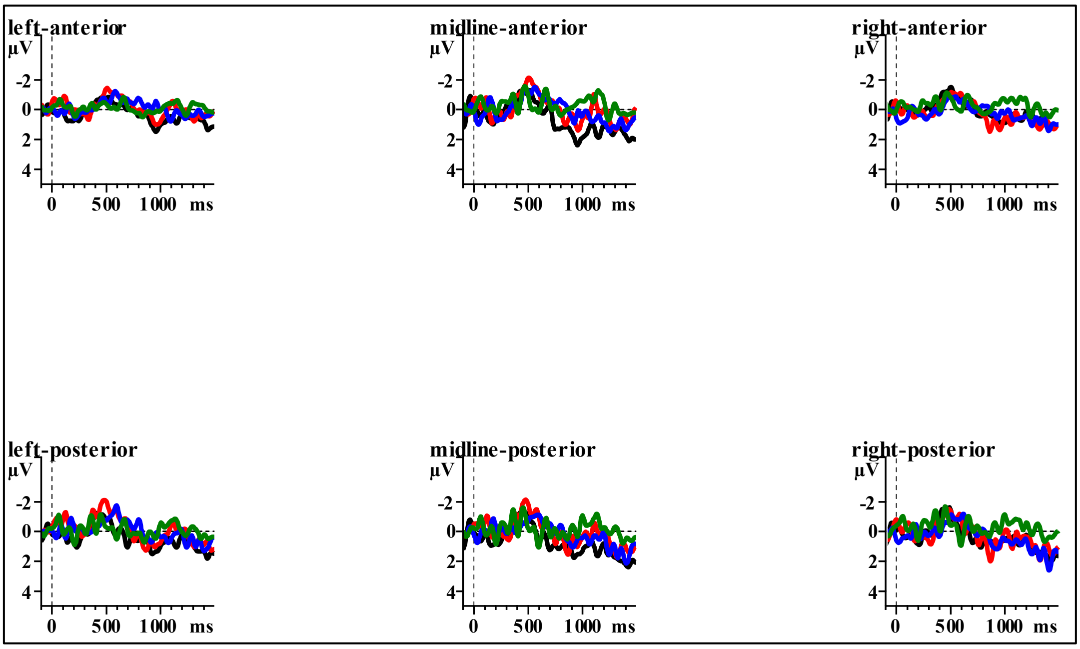
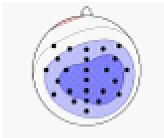
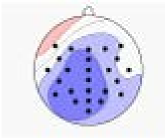
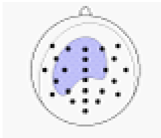
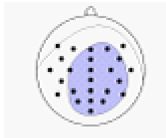


Figure 4.10: Difference wave forms of grand average ERPs time-locked to the critical direct object NP according to the Congruity conditions displayed for four groups separated only for visual purpose: black = native speakers, red = EAHP, blue = LAHP and green = LALP. Voltages are plotted on y-axis ranging from -5μV to +5μV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. ROIs are labelled accordingly. Negative voltages are plotted up.

[73] The distinction between early and late AoA is adopted from Meisel (2011, pp. 202) and applied for visual purpose only. The present investigations treat AoA as a continuous factor.

[74] The distinction between high and low proficiency infers the categorization in accordance with the CEFR—i.e., advanced and intermediate, respectively (see Chapter 1). This categorization, again, is applied for visual purpose only. The present investigations treat proficiency as a continuous factor.

Native Speakers AoA = 0 proficiency > 80	EAHP AoA < 8 proficiency > 80	LAHP AoA ≥ 8 proficiency > 80	LALP AoA ≥ 8 proficiency ≤ 80
 400 ms - 500 ms	 400 ms - 500 ms	 400 ms - 500 ms	 400 ms - 500 ms

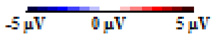


Figure 4.11: Difference voltage maps illustrating the mean differences of grand average ERPs time-locked to the critical direct object NP according to the Congruity conditions in the time-window 400–500 ms post stimulus. Differences are displayed for four groups separated only for visual purpose. Groups are indexed at the upper row of each column. Time range representing the average voltage difference for each head is labelled accordingly. Difference voltage range is plotted from -5µV (dark blue) to +5µV (dark red).

Visual inspection of Figure 4.9 and Figure 4.10 reveals an enhanced negativity for native speakers and all L2 learner groups. This negativity appears strongest between 400 and 500 ms. Differences across groups superficially appear with regard to the distribution rather than the strength and latency of the negativity. Native speakers' negativity appears the strongest on central scalp sites with a right-lateral bias. EAHP shows largest negativities on central and left-lateralized ROIs. For LAHP, the negativity seems to be more frontally distributed, while LALP shows a rather central negativity. Latency differences are not clearly observable. Statistical analyses are carried out for the time-window between 400 and 500 ms, where the negativity seems the strongest for all groups. The time-window is analysed separately for native speakers and L2 learners (see below). See Chapter 3.4.4 for a detailed description of the mixed-effects models. AoA and proficiency did not enter the model performed to analyse the native speaker's ERP data.

4.2.3.1 Time-Window: 400–500 ms

Statistical results of native speakers' ERPs reveal a main effect for ROI (F(5,758) = 32.11, p < .001)[75]. The mean potentials are more negative on anterior than posterior ROIs. Visual inspection indicates an effect for Congruity; however, this impression fails to reach statistical significance. The original (and most complex) statistical model shows a marginal trend for Congruity (F(1,780) = 3.28, p = .07)[76], but it lacks statistical reliability. Moreover, the interaction with ROI does not yield reliability either (F(5,780) < 1)[77]. Although Congruity and its interaction as a fixed effect were removed from the model, the *post hoc* analyses of separate ROIs were performed due to planned comparisons. Single ROI analyses reveal a significant effect for Congruity only in the right-posterior ROI (F(1,137) = 5.05, p = .03)[78]. The detailed statistical analyses of single ROIs are given in appendix 2. Despite the fact that this is not a very sufficient result, it still approaches native-like processing of semantic incongruity and will be seen as an instance of the N400 effect. Possible reasons for the statistical absence of a broader Congruity effect in native speakers' ERPs are discussed in Chapter 4.2.4 below.

Table 4-3: ANOVA table of the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners according to the Congruity conditions (*denominator upper bound df* = 2257, *denominator lower bound df* = 2077).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CONGRUITY	1	8.00	8.00	8.42	< .001	< .001
ROI	1	148.28	29.66	30.55	< .001	< .001
AOA	1	0.93	0.93	< 1		
CONGRUITY×ROI	5	16.51	3.3	3.4	< .001	< .001
ROI×AOA	5	17.74	3.55	3.65	< .001	< .001

The statistical results for L2 learners' ERPs are displayed in Table 4-3. A global analysis reveals a significant main effect for Congruity (-0.84μV), indicating that the mean voltage for the incongruent condition (-0.47μV) is more negative on average than for

[75] *Denominator lower bound* = 596, p < .001.
[76] *Denominator lower bound* = 636, p < .001.
[77] *Denominator lower bound* = 636, p = .32.
[78] *Denominator lower bound* = 101, p = .03.

the congruent condition (0.37μV). The highly significant two-way interaction indicating distributional differences for the strength of the Congruity effect is resolved by ROIs. The results of *post hoc* single ROI analyses are given in Table 4-4.

Table 4-4: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners. Diff = Difference.

ROI	*denominators upper-/lower- bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	462/342	-0.57	1	5.48	5.48	4.33	.04	.04
MIDLINE-ANTERIOR	352/232	-1.13	1	4.14	4.14	9.67	< .001	< .001
RIGHT-ANTERIOR	469/349	-0.69	1	5.16	5.16	6.77	.01	.01
LEFT-POSTERIOR	469/349	-0.97	1	6.63	6.63	7.72	.01	.01
MIDLINE-POSTERIOR	351/231	-1.01	1	5.27	5.27	6.75	.01	.01
RIGHT-POSTERIOR	462/342	-0.83	1	6.64	6.64	7.48	.01	.01

The congruity effect is reliable on all ROIs and appears to be numerically the largest on midline-anterior electrodes. The main effects for AoA and proficiency do not reach reliability. Furthermore, they do not yield any significant influence on the strength or distribution of the Congruity effect. While the factor proficiency is completely removed from the model, increasing AoA shows influence on the distributional differences of the mean potentials: They are more negative-going on anterior than posterior ROIs. Again, this does not impact Congruity. An AoA impact is reliably found only in the left-anterior ROI ($F(1,462) = 4.52$, $p = .03$)[79], revealing that with increasing AoA the mean potentials are gradually going more negative .

The average peak latency of the difference waves within this time window (400–500 ms) was determined for each participant and analysed in a linear regression analysis, including AoA and proficiency as independent factors. The results, however, do not reveal any significant effects except for the intercept.

[79] *Denominator lower bound* = 342, $p = .03$.

4.2.4 Summary and Discussion: Semantic Incongruity

As was expected, overall ERPs reveal an enhanced negativity for the incongruent condition compared with the congruent condition. With respect to the linguistic material, this will be seen as an instance of the N400 effect; reflecting differences of the processing difficulties associated with semantic integration mechanisms (as discussed in Chapter 2.2.1.1). The global analysis lacks statistical significance of the N400 effect for German native speakers, even though the visual inspection of both Figure 4.9 and Figure 4.10 infers an enhanced negativity for the incongruent condition. Additionally, single ROI analysis reveals a significant negativity effect on right-posterior scalp-sites, which corresponds to the classic distribution of the N400 effect. There might be several reasons for the lack of a broad N400 effect: For instance, it could be the case that variance between data points is far enhanced and therefore the amount of data might be too little in order to show a clear N400 effect. This is a rather common observation in psycholinguistic experimental research, especially when more complex statistics such as mixed-effects models are performed. Such models usually require a solid amount of data or high power studies (S. Vasishth, personal communication, September 10, 2013, see also Vasishth's own comments on his blog http://vasishth-statistics.blogspot.de/2014/11/should-we-fit-maximal-linear-mixed.html [last retrieved on August 6, 2015]). However, with respect to enhanced variation of the data and the fact that this might hardly be explainable by the present model there weren't any problems with the convergence of the model. Therefore, the statistical power of the present data elicited by native speakers is considered sufficient. A second possible explanation for the failure of the N400 effect to reach significance might be the choice of the time-window. In order to investigate this, four smaller time-windows (range 25 ms) within the 400–500 ms range were cut and analysed. Entire statistical results are given in appendix 2. Within a shorter time-window between 475–500 ms native speakers reveal a significant N400 effect with a right-posterior distribution (for the corresponding statistical results see appendix 2), which is commonly found in the monolingual processing of semantic incongruity (see Chapter 2.2.1.1). In the remaining time-windows neither the negativity nor distributional differences in its strength reach statistical significance. These results resemble those of the broad time-window reported above. However, the planned *post hoc* analyses of single ROIs constantly yield reliable negativity effects on right-lateral ROIs relative to the remaining central and left-lateral ROIs. As for now, it is assumed that native German speakers are sensitive when reading sentences, including a semantic incongruity. In the present data this sensitivity is reflected by a late and rather short negativity effect that is globally strongest between 475 and 500 ms post stimulus presentation, showing a right-posterior distribution. Within the broader time-window, the negativity effect is consistently found on right-posterior electrodes. Hence, it will be seen as an instance of N400. One further observation that may possibly account for a weakening of the present N400 effect may be made with respect to the stimulus material. As can be seen in appendix 5.1 some critical items contain a rather weak semantic violation that becomes strong only until the end of the sentence (thanks to Sophie Repp for this insightful comment). Nonetheless, the majority of the sentences used as stimulus material include a strong semantic violation on the critical item. Accordingly, in light of the integrative view on the N400 effect, it presently is acknowledged as an indicator of higher processing costs due to the enhanced activation and retrieval of neural mechanisms that are responsible for updating the current context with new information.

Results of L2 learners' ERPs suggest that in general semantic processing mechanisms are activated and retrieved in L2 processing. This means that processing mechanisms responsible for updating the prior context with the new information reflect difficulties when the incoming information is not or less predicted. Results of the smaller time-windows, which were also analysed for L2 learners—in order to remain comparability with data obtained for native speakers—reveal similar results as stated above with respect to the time-windows between 450–475 ms and 475–500 ms. The corresponding tables are given in appendix 2. Relative to the outcomes of former studies and the hypotheses proposed in Chapter 4.2.1, the present results may be interpreted in two ways: On the one hand, the results largely confirm the overall understanding of semantic L2 processing so that L2 learners robustly activate and retrieve the relevant neural resources and mechanisms—i.e., L2 learners elicit an N400 effect. On the other hand, and with respect to the impact of the AoA and proficiency on L2 processing, the present results differ from those of former investigations and previous expectations. Former studies revealed influences of either AoA or proficiency on the N400 effect relative to strength, distribution, and latency (again, see Chapter 2.2.1.2). Yet, neither of the impacts is present in the current results. There are reliable distributional differences in the strength of the N400 effect; however, those are not predicted by either AoA or proficiency influence. Moreover, strength and latency changes are not observed. Previous results reported frontally reduced amplitudes of the N400 effect as well as delayed latencies that were preponderantly accounted to proficiency influence (Ardal et al., 1990; Newman et al., 2012; Ojima et al., 2005). Nonetheless, according to the design of the present statistical model, the current results suggest that L2 proficiency is not predictive at all with respect to changes in the strength, latency, and also distribution of the N400 effect. Regarding the distribution of the mean potentials elicited by L2 learners increasing AoA generally yields more negative-going mean potentials, especially on the left-anterior ROIs. Observations of more left-lateralized negativities have also been made by former studies (Ardal, et al., 1999; Ojima et al., 2005; Weber-Fox and Neville, 1996). Yet, these results directly accounted for the distribution of the N400 effect—i.e., for the difference of average ERPs across conditions, not for the mean potentials in general.

In summary, the present results suggest that the L2 processing of semantic incongruity is robust and similar to that of semantic L1 processing. AoA and proficiency influence on the N400 effect is not reported. Thus, given that the overall proficiency of L2 learners is relatively high (see Chapter 4.1.1), it doesn't seem to have any impact on semantic L2 processing. In other words, present data suggests that when a solid and hence intermediate proficiency level is reached, L2 learners consistently anticipate the semantic incongruities. This is reflected by the robust activation of the neural resources associated with semantic processing mechanisms, irrespective of AoA. Consequently, the question whether AoA and / or proficiency have influence over the semantic processing of L2 and whether the influences appear to be gradually cannot be answered on the basis of the current results.

4.3 Experiment 2: Double Nominative Violation

4.3.1 Stimulus Materials and Hypotheses

In the second experiment, ERPs as response to the L1 and L2 processing of German double nominative violation are investigated. Sample stimulus sentences are presented in Table 4-5; see also appendix 5.2 for the full list of stimulus sentences. The item sentences include a finite transitive verb, preceded by a subject NP with nominative case-marking expressed by the definite determiner 'der' ('the$_{\mathbf{NOM}}$'). All subject NPs are masculine, singular, and animate. The second NP following the transitive verb and prior to the coordination 'und' ('and') serves as the critical item. This second NP either is correctly case-marked by the masculine singular accusative and definite determiner 'den' ('the$_{\mathbf{ACC}}$') or incorrectly case-marked by the masculine singular nominative determiner 'der' ('the$_{\mathbf{NOM}}$'). The latter indicates the double nominative violation, because German does not allow identical case-marking within one single clause.[80] In addition, all critical item NPs are masculine, singular, and non-animate.[81]

Table 4-5: Stimulus examples of case-congruent and case-violation sentences: NOM = Nominative, ACC = Accusative, '*'-indexed sentences entail a case-violation.

<u>Correct NOM-ACC condition (case-congruent)</u>
Der$_{\mathrm{NOM}}$ Sportler$_{\mathrm{NOM}}$ gewinnt *den*$_{\mathrm{ACC}}$ *Wettkampf*$_{\mathrm{ACC}}$ und erhält einen Pokal$_{\mathrm{ACC}}$
The$_{\mathrm{NOM}}$ *athlete*$_{\mathrm{NOM}}$ *wins the*$_{\mathrm{ACC}}$ *competition*$_{\mathrm{ACC}}$ *and gets a trophy*$_{\mathrm{ACC}}$.
<u>Incorrect NOM-NOM condition (case-violation)</u>
*Der$_{\mathrm{NOM}}$ Sportler$_{\mathrm{NOM}}$ gewinnt *der*$_{\mathrm{NOM}}$ *Wettkampf*$_{\mathrm{NOM}}$und erhält einen Pokal$_{\mathrm{ACC}}$.
**The*$_{\mathrm{NOM}}$ *athlete*$_{\mathrm{NOM}}$ *wins the*$_{\mathrm{NOM}}$ *competition*$_{\mathrm{NOM}}$ *and gets a trophy*$_{\mathrm{ACC}}$.

[80] Similar to the stimulus material used for Experiment 1 (see Chapter 4.2.1), two additional filler conditions were included. Both filler sentences started with a subject QP 'Kein Sportler' ('no athlete') and were included to avoid reading strategies by the participants. See also Footnote 62 for further explanations.

[81] Polish morphologically marks accusative masculine singular animate nouns by '–a', for exceptions see e.g., Engel, Rytel-Kuc, Cirko, and Debski (1999, pp. 762). As for non-animate masculine singular nouns, no morphological suffix marking is required in neither language, which should minimize the possible L1-(Polish) influence for L2 processing.

The linear word order of the present sentences resembles the structure used for stimulus sentences in Experiment 1 (Chapter 4.2.1). Again, the first part of the sentence, including the critical item, has an SVO word order which corresponds to the unmarked / canonical word order of the matrix sentence structure in both languages German and Polish (but see Footnote 23). It includes a nominative case-marked subject NP typically bearing agent role[82], followed by a finite transitive verb which is inflected in agreement with the preceding nominative case-marked subject NP. The verb precedes the direct object NP with an accusative case-marking, typically bearing a patient or theme role.[83] Again, in order to avoid sentence wrap-up effects (see Footnote 63), each sentence is followed by the conjunction 'und' ('and'), a finite transitive verb (inflected in agreement with the subject NP) (see also Chapter 4.2.1) and terminate with a direct object NP case-marked with accusative (all masculine, singular, and non-animate). Note that German and Polish differ in their morphological case-marking of nominative / accusative with respect to the presence or absence of the appropriate definite determiner[84], respectively.

According to the findings of former ERP studies on the German L1 processing of double nominative violations (see Chapter 2.2.2.2) an enhanced P600 component on the nominative case-marked critical item should be elicited by native speakers' ERPs. More precisely, the processor starts off with the processing of the first NP by assigning a default subject role to it in accordance with its categorical and case-features. Next, the processor comes across the transitive verb, which agrees with the preceding NP in all appropriate features, thereby strengthening the former prediction that the initial NP ought to be the subject to the present sentence. If the transitive verb is followed by an accusative case-marked NP, the processor will not have any difficulties to syntactically integrate the noun as direct object into the present sentence structure and thereby assigning the patient role. However, when the processor encounters a nominative case-

---

[82] The assignment of agent-role to the subject NP should be further emphasized due to its animacy marking.
[83] Thematic role assignment is further stressed by the non-animacy of the critical items; see also Footnote 82 above.
[84] In Polish definiteness is marked rather via word order and not through determiners.

marked NP after the verb, it is expected to face difficulties that are due to a structural mismatch with the former prediction. In other words, as the first NP is case-marked nominative and the following verb does not violate subject-verb agreement rules due to the canonical word order, the language processing system will hardly ever expect another nominative case-marked NP but rather an accusative case-marked NP. This failure of correct (and predicted) case-marking on the second NP then should trigger syntactic processing costs which are reflected by the P600. This P600 is expected to mirror enhanced activation and retrieval of neural mechanisms associated with syntactic repair since the structure is ungrammatical due to case-marking failure. Enhanced processing costs due to thematic hierarchizing (Frisch & Schlesewsky, 2001) are not expected: for the present stimuli sentences, processing demands inferring thematic hierarchizing should not enhance due to the respective animacy marking of both NPs. To the processing system, it should be clear that the first nominative case-marked (and animate) NP is assigned the role of an agent, while the second non-animate NP is assigned the role of a patient, irrespective of correct or incorrect case-marking.

With respect to the former findings of ERP patterns as response to the L2 processing of double nominative violations, they are expected to be different compared with that of native speakers. Recollect from Chapter 2.2.2.4 that the two main objectives are (i) whether and how syntactic L2 processing patterns differentiate from those of L1; and (ii) whether changes in syntactic L2 processing development appear gradually in general due to either improving proficiency, or increasing AoA, or both. Given the correlation of AoA and proficiency reported in Chapter 4.1.1, it is of particular interest whether and how the L2 processing patterns are impacted by these two factors. Accordingly, recent results suggest that changes in syntactic L2 processing mechanisms are determined by an interaction between AoA and proficiency influence (see again Chapter 2.2.2.3). At late AoA, syntactic processing mechanisms are claimed to be intact, but less controlled and only elicit the appropriate ERP component (e.g., P600) provided that a proficiency has been reached to an appropriate level (Steinhauer et al., 2009; see also McLaughlin et al., 2010; Osterhout et al., 2006; Tanner et al., 2013). A weaker P600 then is ought to mirror this reduced control on the activation of syntactic processing mechanisms. With respect to the continuous character of AoA and proficiency, it might further be assumed that, on the one hand, increasing AoA gradually reduces the strength of the P600. On the other hand, improving proficiency might enhance the P600. Until now, whether there are differences in the weighting of these two influences and whether data analysis may account for the suggested interaction between AoA and proficiency, and also whether the potential influences appear gradually rather than discontinuous, remains uncertain. The next sections report the behavioural (Chapter 4.3.2) and ERP results (Chapter 4.3.3) which are further discussed in Chapter 4.3.4.

4.3.2 Behavioural Results: Double Nominative Violation

Table 4-6 summarizes the average accuracy ratings obtained for naïve speakers and L2 learners. Figure 4.11 displays the mean accuracy ratings obtained for each participant. Results of the statistical analysis of the accuracy rates for native speakers do not show any rating differences between case-congruent sentences and sentences including a case-violation (model only had an intercept).

Table 4-6: Mean accuracy rates (in %) and RTs (in ms) revealed by native speakers and L2 learners for case-congruent and case-violation conditions with standard deviations in parentheses.

Group	**Accuracy in %**		**Reaction Times in ms**	
	case-congruent	case-violation	case-congruent	case-violation
Native Speakers (n = 18)	94.02 (23.72)	93.08 (25.39)	454.86 (234.02)	438.33 (186.84)
L2 learners (n = 60)	80.08 (39.95)	64.27 (47.93)	656.54 (418.98)	576.19 (386.74)

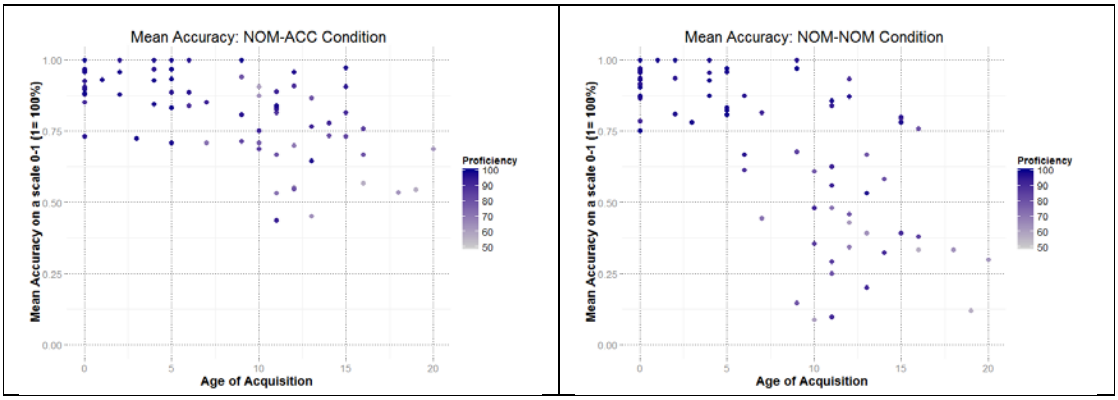


Figure 4.12: Mean accuracy rates (in %) for both, case-congruent (NOM-ACC, left) and case-violation (NOM-NOM, right) conditions for all participants (n = 78). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, accuracy = 100 % and shading = 90-100) are mapped only once.

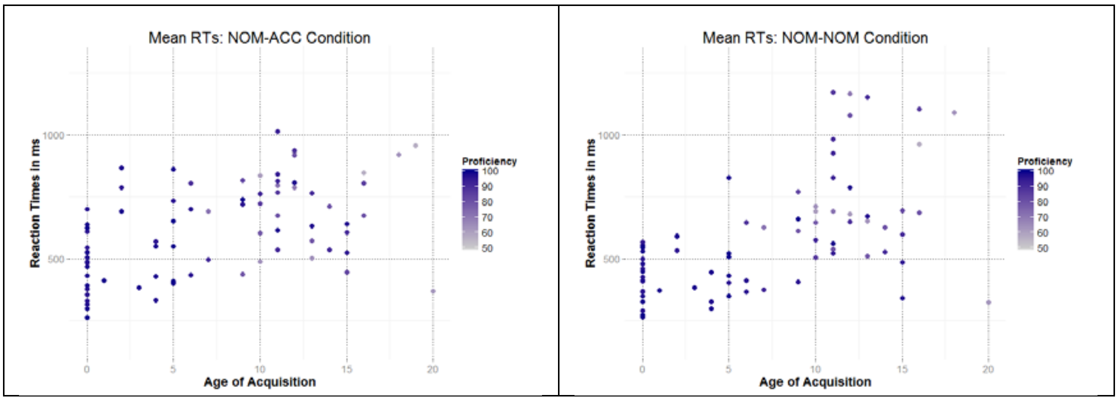


Figure 4.13: Mean reaction times (in ms) for both case-congruent (NOM-ACC, left) and case-violation (NOM-NOM, right) conditions for all participants (n = 78). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, RT = 500 ms and shading = 90-100) are mapped only once.

The statistical analysis of L2 learners' ratings reveals a significant main effect for Case ($F(1,3389) = 17.5$, $p < .001$)[85]. Generally, case-congruent sentences (NOM-ACC) are rated more accurately than case-violation sentences (NOM-NOM). Ratings significantly become more accurate as proficiency improves ($F(1,3389) = 67.77$, $p < .001$)[86]. Further, this proficiency impact is reliably more pronounced for the ratings of the case-congruent condition ($F(1,3389) = 7.77$, $p = .005$)[87]. Additionally, AoA influence is

[85] *Denominator lower bound* = 3179, $p < .001$.
[86] *Denominator lower bound* = 3179, $p = .005$.
[87] *Denominator lower bound* = 3179, $p < .001$.

predictive as to accuracy ratings ($F(1,3389) = 11.9, p < .001$) [88] in such a manner that the increasing AoA generally decreases accuracy. An AoA impact on accuracy differences across conditions is not reported.

RT outcomes for only accurately rated sentences are displayed in Figure 4.12. Groups' averages are given in Table 4-6.[89] The analysis of RT data carried out for native speakers does not reveal any significant results (model only had an intercept). Hence, they are equally fast in accurately judging a case-congruent as well as a case-violation sentence. The results of the statistical analysis of L2 learners' RTs reveal a main effect for case ($F(1,2410) = 3.89, p = .04$) [90]. This indicates that on average judgements for case-congruent sentences (NOM-ACC) are performed significantly faster than for case-violation sentences (NOM-NOM). Moreover, improving proficiency reliably yields the speeding up of RTs for sentences including a case-violation (NOM-NOM), whereas the speed to judge a case-congruent sentence is not as much affected ($F(1,2410) = 4.35, p = .03$) [91]. Further, Figure 4.12 shows that increasing AoA slows down the RTs to judge a case-violation condition ($F(1,2410) = 10.31, p = .001$) [92].

In short, behavioural results suggest that relative to the case-violation condition (NOM-NOM), improving proficiency increases the accuracy and the speed of sentence judgements, whereas increasing AoA yields alleviation of both, accuracy and speed of RTs. Though, results of the behavioural performance are far removed from the critical items; and accuracy judgements and RTs also could be influenced by sentence completions: Therefore, these behavioural outcomes will not be discussed any further.

---

[88] *Denominator lower bound* = 3179, $p < .001$.
[89] Accuracy rates elicited by L2 learners indicate that—especially for the case-violation condition—a great data loss (approx. 35%) has to be considered when removing all inaccurately rated sentences. Hence, RT data were additionally analysed including all data points in the model: This additional analysis reveals similar results (i.e. effect for case ($F(1,3333) = 16.77, p < .001$)) except for the lack of the significant interaction between Case and proficiency ($F(1,3333) = 2.34, p = .12$). Yet again, increasing AoA reveals slower judgments for the case-violation condition (NOM-NOM) than for the case-congruent condition. RTs when judging the case-congruent condition (NOM-ACC), ($F(1,3333) = 10.46, p = .001$). Statistics are presented for the denominator upper bound.
[90] *Denominator lower bound* = 2200, $p = .04$.
[91] *Denominator lower bound* = 2200, $p = .03$.
[92] *Denominator lower bound* = 2200, $p = .001$.

4.3.3 ERPs: Double Nominative Violation

ERPs of the double nominative violation included 4,541 trials. Due to artefact and drift-rejection, 593 trials (13.06%) had to be removed from averaging. Figure 4.14 displays the difference waves in a time-window from -100 to 1,500 ms, relative to the critical items on six ROIs.[93] Again, L2 learners were separated into three groups for visual purposes (cf. Chapter 4.2.3). The grand average ERPs illustrating single group comparisons between case-congruent and case-violation conditions are listed in appendix 1.2.

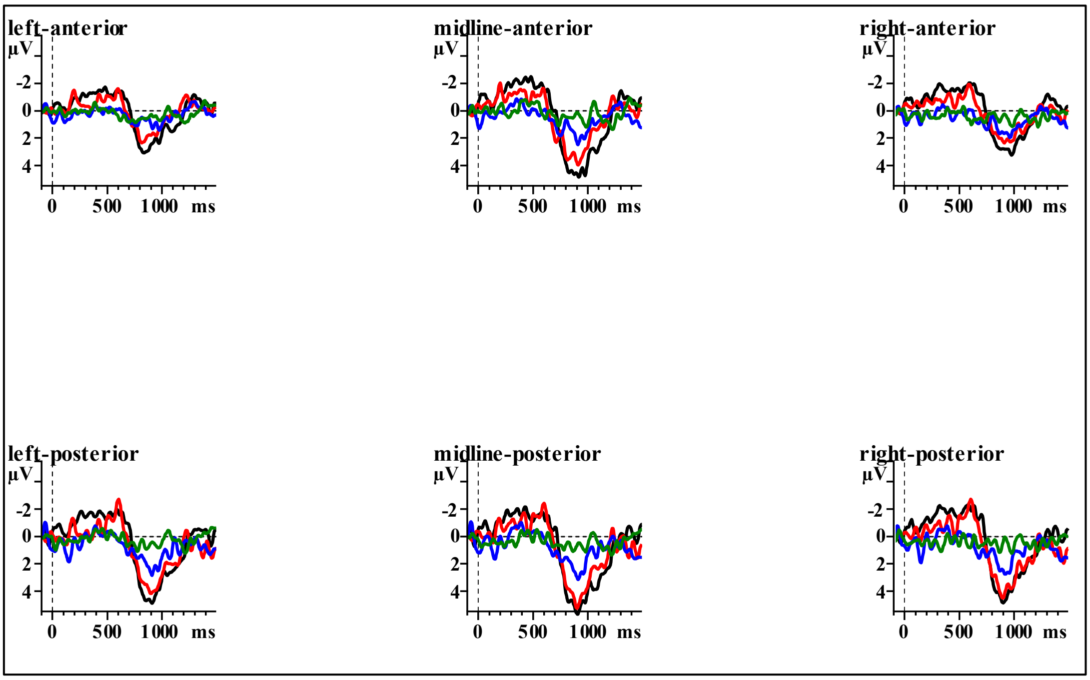


Figure 4.14: Difference wave forms of grand average ERPs time-locked to the critical direct object NPs according to the Case conditions by four groups separated only for visual purpose: black = native speakers, red = EAHP, blue = LAHP and green = LALP. Voltages are plotted on y-axis ranging from -5µV to +5µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. ROIs are labelled accordingly. Negative voltages are plotted up.

[93] Grand average difference was calculated on the basis of the mean voltage of the case-congruent condition subtracted from the mean voltage of the case-violation condition.

Native Speakers AoA = 0 proficiency > 80	EAHP AoA < 8 proficiency > 80	LAHP AoA ≥ 8 proficiency > 80	LALP AoA ≥ 8 proficiency ≤ 80

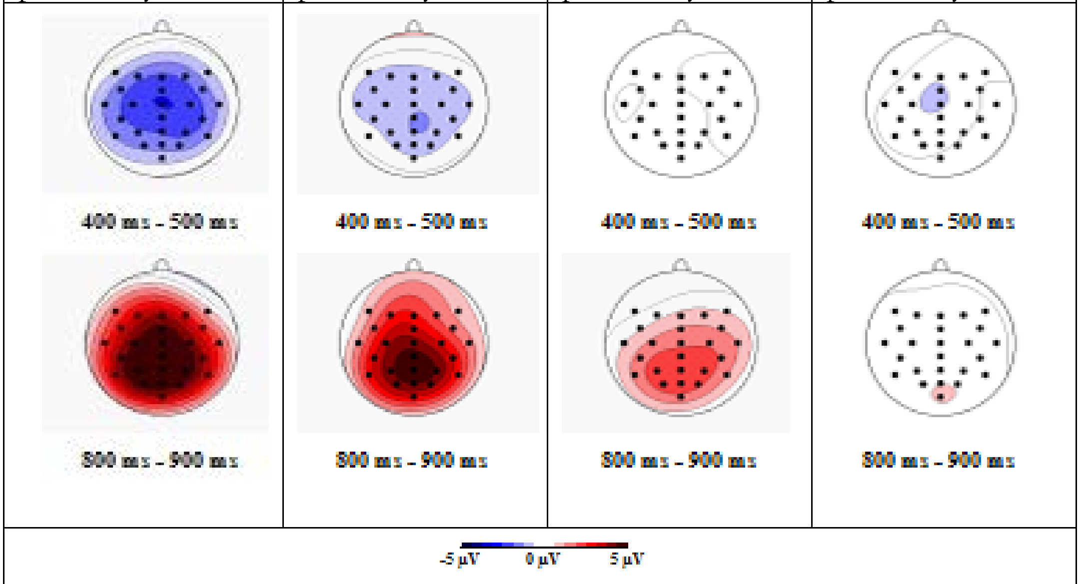


Figure 4.15: Difference voltage maps illustrating the mean differences of ERPs time-locked to the critical direct object NPs according to the Case conditions in two time-windows: 400–500 ms and 800–900 ms post stimulus. Differences are displayed for four groups separated only for visual purpose. Groups are indexed at the upper edge of each row. Time range representing the average voltage difference for each head is labelled accordingly. Difference voltage range is plotted from -5µV (dark blue) to +5µV (dark red).

Figure 4.15 above displays the voltage maps of the difference waves for these two time windows for all four groups. Visual inspection generally suggests differences in the processing patterns between native speakers and L2 learners with different AoA and proficiency levels. Interestingly, native speakers reveal a large early negativity with a central distribution. This negativity is rather unexpected relative to hypotheses that were formulated based on former results (see Chapter 4.3.1 above, see also Chapter 2.2.2.4). The early negativity is followed by a large late positivity with a maximum approx. between 800 and 900 ms that appears most prominent on central-posterior electrodes. Relative to visual inspection of native speakers, two time-windows were cut for further analyses: the early one 400–500 ms and the late one 800–900 ms.

Accordingly, both components the early negativity and the late positivity appear stronger for the native speakers' ERPs. Regarding L2 learners' ERPs the negativity seems to continuously reduce with respect to increasing AoA. The late positivity, also, seems to attenuate with increasing AoA, although this alleviation seems less striking. Further, the visual attenuation of the positivity appears more pronounced due to proficiency influence. Separate analyses for native speakers and L2 learners of the two time-windows are presented in turn.

4.3.3.1 Early Time-Window: 400–500 ms

Iterative calculation of the linear mixed-effects model carried out for native speakers' ERPs in the 400–500 ms time-window reveal highly significant main effects for Case ($F(1,759) = 14.53$, $p < .001$)[94] and ROI ($F(5,759) = 3.69$, $p < .001$)[95]. The average potential is more negative in the case-violation (-1.73μV) than in the case-congruent condition (0.03μV). This difference is equally distributed over the entire scalp although left-lateral electrodes reveal more positive mean potentials than midline- and right-lateral electrodes (see appendix 3 for statistical results on planned single ROI analyses). With respect to the stimulus material and time-window this negativity may be seen as an instance of the N400 effect. This finding is rather unexpected and will be returned to in the discussion in Chapter 4.3.4.

Final results of the statistical analysis of L2 learners' ERPs in the 400–500 ms time-window are listed in Table 4-7. Although the mean potentials of the case-violation (-0.06μV) are more negative than those of the case-congruent condition (0.26μV), this difference does not reach statistical significance. The two-way interaction between case and ROI indicates that the mean differences across conditions vary as to scalp distribution and hence is resolved by ROI. Yet, as shown in Table 4-8, the *post hoc* analyses of single ROIs do not yield any significant negativity effects. Numerically, the

[94] *Denominator lower bound* = 615, $p < .001$.
[95] *Denominator lower bound* = 615, $p < .001$.

negativity is the largest on midline ROIs. With respect to laterality, it appears larger on left-lateral than right-lateral ROIs. Topographical differences in terms of anteriority are not observed.

Table 4-7: ANOVA table of the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners according to the Case conditions (*denominator upper bound df* = 2574, *lower bound df* = 2094).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CASE	1	1.89	1.89	1.91	.16	.16
ROI	5	181.01	36.2	36.7	< .001	< .001
AOA	1	2.14	2.14	2.17	.15	.15
CASE×ROI	5	12.62	2.52	2.56	.05	.05
CASE×AOA	1	3.93	3.92	4.00	.05	.05

Table 4-8: ANOVA table of the Case effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower- bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	478/358	-0.35	1	3.42	3.42	2.17	.14	.14
MIDLINE-ANTERIOR*	358/238	-0.55	1	1.41	1.41	2.82	.09	.09
RIGHT-ANTERIOR*	478/358	-0.13	1	0.29	0.29	< 1		
LEFT-POSTERIOR*	478/358	-0.34	1	1.48	1.48	1.4	.23	.23
MIDLINE-POSTERIOR*	358/238	-0.53	1	1.74	1.74	2.12	.14	.14
RIGHT-POSTERIOR*	478/358	-0.14	1	0.26	0.26	< 1		

Results further show that the strength of the negativity is subject to AoA influence. Figure 4.16 illustrates that with increasing AoA, the strength of the negativity continuously attenuates due to slightly positive-going potentials of the case-violation condition and more pronounced negative-shifting potentials of the case-congruent condition. Separate analyses of AoA influence on each case condition do not yield statistical significance. Finally, proficiency was removed from the model, which indicates that it does not affect the mean potentials in the early time-window.

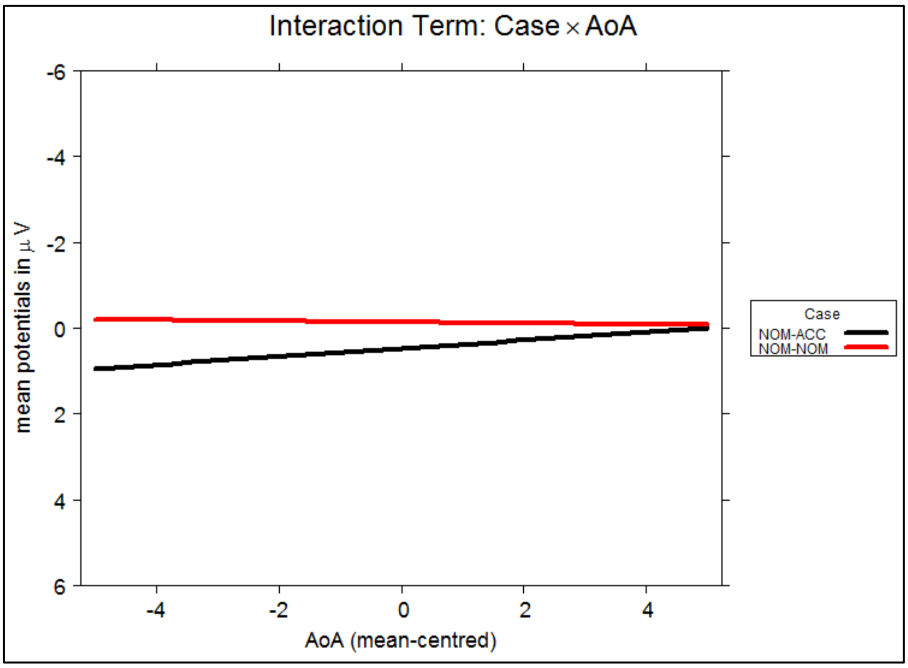


Figure 4.16: Plot of Interaction Term "Case × AoA" according to the time-window 400–500 ms. AoA data (mean-centred) is mapped on x-axis. Y-axis represents the mean voltages (range from 6μV to -6μV) of each condition: case-congruent (NOM-ACC, black line) and case-violation (NOM-NOM, red line). Negative voltages are plotted up.

4.3.3.2 Late Time-Window: 800–900 ms

Statistical analyses of the ERPs elicited by native speakers in the late time-window reveal a significant effect for Case ($F(1,759) = 37.76$, $p < .001$)[96], indicating a reliable positivity effect. Furthermore, results show a significant effect for ROI ($F(1,759) = 19.28$, $p < .001$)[97] and the reliable two-way interaction with Case ($F(1,759) = 14.45$, $p < .001$)[98]. The positivity effect (average voltage difference = 3.91μV) is numerically larger on posterior than anterior ROIs. Corresponding *post hoc* analyses of single ROIs are summarized in Table 4-9 below. With respect to the stimulus material and expected processing pattern (see Chapter 4.3.1), this positivity effect will be seen as an instance of P600.

[96] *Denominator lower bound* = 615, $p < .001$.
[97] *Denominator lower bound* = 615, $p < .001$.
[98] *Denominator lower bound* = 615, $p < .001$.

Table 4-9: ANOVA table of the Case effect in each ROI according to the average ERP amplitudes in the time-window 800–900 ms conducted for native speakers. Diff = Difference.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	139/103	3.08	1	44.39	44.39	31.1	<.001	<.001
MIDLINE-ANTERIOR	106/70	4.43	1	53.35	53.35	37.89	<.001	<.001
RIGHT-ANTERIOR	140/104	2.54	1	46.72	46.72	17.8	<.001	<.001
LEFT-POSTERIOR	140/104	4.41	1	102.77	102.77	37.31	<.001	<.001
MIDLINE-POSTERIOR	102/66	5.35	1	50.01	50.01	45.47	<.001	<.001
RIGHT-POSTERIOR	140/104	4.09	1	83.49	83.49	32.1	<.001	<.001

Table 4-10: ANOVA table of the average ERP amplitudes in the time-window 800–900 ms conducted for L2 learners according to the Case conditions (*denominator upper bound df* = 2538, *denominator lower bound df* = 2058).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CASE	1	69.59	69.56	49.28	< .001	< .001
ROI	5	189.63	37.93	26.87	< .001	< .001
PROFICIENCY	1	3.95	3.95	2.8	< .1	< .1
AOA	1	2.22	2.22	1.57	< .3	< .3
CASE×ROI	5	165.26	33.05	23.41	< .001	< .001
CASE×PROFICIENCY	1	22.01	22.01	15.59	< .001	< .001
CASE×AOA	1	0.51	0.51	< 1		
ROI×PROFICIENCY	5	48.97	9.79	6.94	< .001	< .001
ROI×AOA	5	5.45	1.09	< 1		
PROFICIENCY×AOA	1	1.66	1.66	1.17	< .3	< .3
CASE×ROI×PROFICIENCY	5	63.64	12.73	9.02	< .001	< .001
CASE×ROI×AOA	5	32.05	6.41	4.54	< .001	< .001
CASE×PROFICIENCY×AOA	1	13.14	13.14	9.31	< .001	< .001

The statistical results of brain responses elicited by L2 learners' ERPs in the late time-window are given in Table 4-10. Differences between the mean potentials across conditions reveal a positivity effect (mean voltage difference = 2.3μV) which on average is smaller than that reported for native speakers. Distributional differences of the strength of the positivity appear statistically significant. The results of the planned *post hoc* analyses are summarized by Table 4-11. Numerically, the positivity is larger on posterior than anterior ROIs. Moreover, distributional differences across conditions are reliably impacted by both AoA and proficiency.

Table 4-11: ANOVA table of the Case effect in each ROI according to the average ERP amplitudes in the time-window 800–900 ms conducted for L2 learners. Diff = Difference.

ROI	*denominators upper-/lower-bound*	*mean Diff in µV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	462/342	1.14	1	21.16	21.16	18.71	<.001	<.001
MIDLINE-ANTERIOR	345/225	1.76	1	19.88	19.88	27.33	<.001	<.001
RIGHT-ANTERIOR	461/341	1.2	1	28.19	28.19	26.56	<.001	<.001
LEFT-POSTERIOR	466/346	2.08	1	84.53	84.53	40.59	<.001	<.001
MIDLINE-POSTERIOR	342/222	2.63	1	20.26	20.26	43.41	<.001	<.001
RIGHT-POSTERIOR	464/344	2.09	1	78.94	78.94	53.28	<.001	<.001

Figure 4.17 illustrates that improving proficiency yields an increase of the positivity effect due to gradually positive-going mean potentials of the case-violation condition. Changes in the mean potentials of the case-congruent condition appear less striking and rather negative-shifting. This proficiency impact is apparent on all ROIs and appears more pronounced on midline and posterior scalp-sites, which is further confirmed by the statistical *post hoc* results given in Table 4-12.

Table 4-12: ANOVA table of proficiency influence on the Case effect in each ROI according to the average ERP amplitudes in the time-window 800–900 ms conducted for L2 learners.

ROI	*Denominators upper-/lower-bound*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	462/342	1	5.86	5.86	5.18	.02	.02
MIDLINE-ANTERIOR	345/225	1	8.89	8.89	12.22	<.001	<.001
RIGHT-ANTERIOR	461/341	1	7.41	7.41	6.98	.01	.01
LEFT-POSTERIOR	466/346	1	24.36	24.36	11.7	<.001	<.001
MIDLINE-POSTERIOR	342/222	1	7.02	7.02	15.05	<.001	<.001
RIGHT-POSTERIOR	464/344	1	22.27	22.27	15.04	<.001	<.001

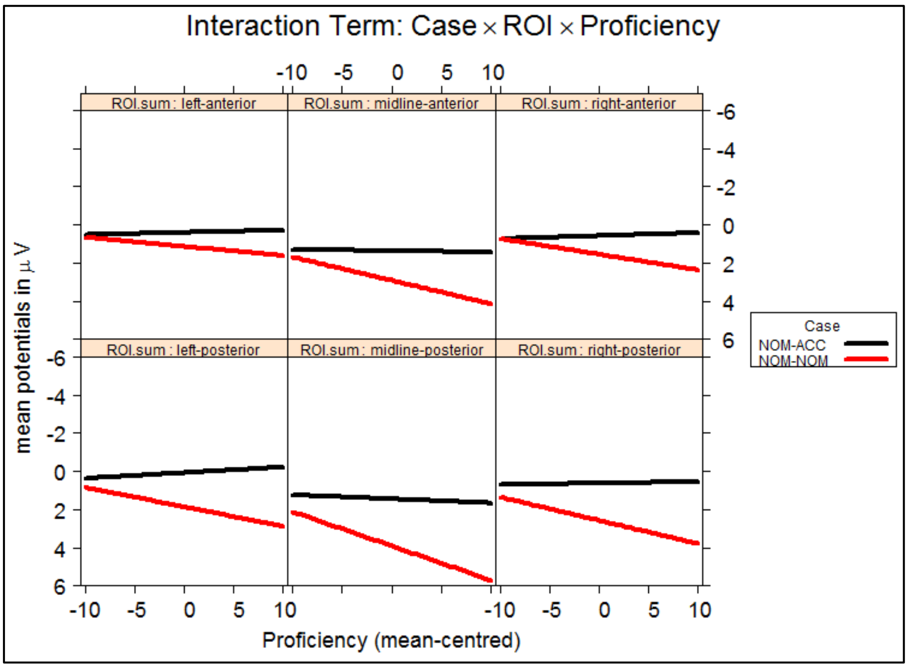


Figure 4.17: Plot of Interaction Term "Case × ROI × Proficiency" according to the time-window 800–900 ms. Proficiency data (mean-centred) is mapped on x-axis. Y-axis represents the mean voltages (range from 6µV to -6µV) of each condition: case-congruent (NOM-ACC, black line) and case-violation (NOM-NOM, red line). Grids represent ROIs (indicated as ROI.sum) and are labelled accordingly. Negative voltages are plotted up.

Figure 4.18 illustrates the influence of AoA on distributional differences in the changing strength of the positivity effect. Superficially, AoA influence is less strong than the proficiency impact (see above). Visibly, the positivity attenuates only on left-anterior scalp-sites due to negative-shifting mean potentials of the case-violation condition. The remaining ROIs do not yield any statistically reliable AoA influence on the case effect. The *post hoc* analyses of single ROIs reveal a significant two-way interaction between Case and AoA only on left-lateral electrode sites ($F(1,462) = 3.91$, $p = .05$)[99] endorsing the visual impression.

[99] *Denominator lower bound* = 342, $p = .05$.

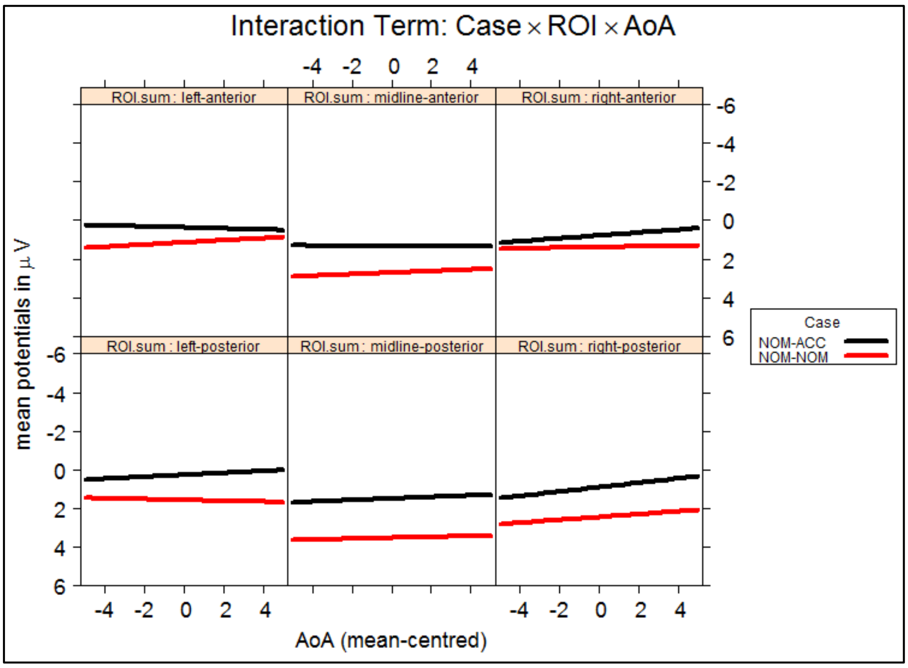


Figure 4.18: Plot of Interaction Term "Case × ROI × AoA" according to the time-window 800–900 ms. AoA data (mean-centred) is mapped on x-axis. Y-axis represents the mean voltages (range from 6µV to -6µV) of each condition: case-congruent (NOM-ACC, black line) and case-violation (NOM-NOM, red line). Grids represent ROIs (indicated as ROI.sum) and are labelled accordingly. Negative voltages are plotted up.

Remarkably, significant changes in the strength of the positivity effect—due to proficiency influence—are further determined by AoA. Visual inspection of Figure 4.19, representing the plot of the term of higher order, suggests that at early AoA the influence of improving proficiency yields a strengthening of the positivity due to continuously positive-going mean potentials of the case-violation condition ($F(1,1284) = 16.19$, $p < .001$)[100]. This influence attenuates as AoA increases. Interestingly, the attenuation of the positivity effect due to decreasing proficiency has impact on increasing AoA; it is also subject to changes in mean potentials of the case-congruent condition. Visually, they appear more negative-going with increasing AoA. Furthermore, with improving proficiency, ERP amplitudes continuously decrease at early stages of AoA. This gradually reverses towards a positive shift of potentials at

[100] *Denominator lower bound* = 1224, $p < .001$.

latest stages of AoA. Statistically, the *post hoc* analysis of single case-congruent conditions does not reveal an interactive influence.

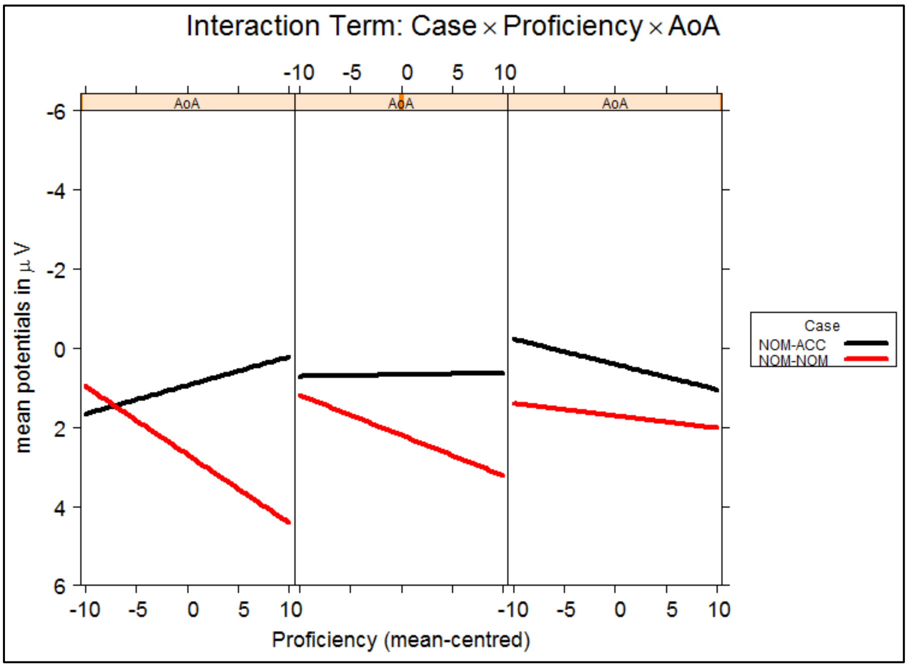


Figure 4.19: Plot of Interaction Term "Case × Proficiency × AoA" according to the time-window 800–900 ms. Proficiency data (mean-centred) is mapped on x-axis. Y-axis represents the mean voltages (range from 6µV to -6µV) of each condition: case-congruent (NOM-ACC, black line) and case-violation (NOM-NOM, red line). Grids display three AoA stages (from mean-centred AoA data): left = earliest AoA, centre = middle AoA and right = latest AoA. Negative voltages are plotted up.

4.3.4 Summary and Discussion: Double Nominative Violation

In the main, the ERP results only partly comply with the hypothesized processing patterns given in Chapter 4.3.1. In accordance with previous expectations outcomes reliably reveal a late positivity effect. As already indicated above, this positivity effect is seen as an instance of P600 reflecting enhanced processing demands due to repair mechanisms, which are activated and retrieved when a double nominative violation occurs. The P600 differs between L1- and L2 processing with respect to strength and distribution. The P600 elicited by L2 learners is less strong compared with that of native speakers' ERPs. It further undergoes continuous changes in its strength which is predominantly due to an interactive influence of AoA and proficiency— i.e., at early AoA, proficiency influence is more exhaustive than at late AoA. A closer look at the proficiency impact on the P600 (elicited by L2 learners) suggests that especially those scalp-regions that usually show the strongest positivity, namely (right-) posterior ROIs, reveal a strengthening of the P600 due to proficiency influence yielding native-like processing. With respect to former findings (see Chapter 2.2.2.2), this largely confirms the idea that improving proficiency triggers native-like activation and retrieval of processing mechanisms, and therefore may compensate for possible AoA influence. This goes in line with the argumentation by Steinhauer et al. (2009) who suggest that the neural resources associated with syntactic repair mechanisms are available as soon as L2 proficiency has reached an appropriate level irrespective of AoA. However, as already indicated, present results also reveal an interactive influence of both factors. Additionally, it is not only the brain responses of the case-violation condition which change through improving proficiency (i.e., attenuating the strength of the positivity due to a negative shift). Also, present data suggests that the way in which a case-congruent structure is processed is subject to proficiency impact which becomes noticeably weaker when AoA is late.[101] Differences regarding the P600 between L1 and L2 processing then are quantitative and also qualitative in nature. For the former, the present results put forward differences in the degree of activation of neural mechanisms. The latter indicates that the resources of the neural mechanisms vary. Additionally, both undergo changes as proficiency improves by yielding more native-like degrees of activation of resources and hence retrieval of the according mechanisms. Further, this proficiency impact is predicted by AoA in such a manner that it is the most exhaustive at early AoA.

Interestingly, prior to the P600, native speakers show a large negativity effect with a global distribution. According to former studies on double nominative violations, this

[101] Accordingly, Hahne (2001) reports differences in the N400 effect between L1 and L2 processing and that these differences are due to the ERPs elicited for semantically congruent rather than semantically incongruent sentences. She argues that the integration of new information into the sentence context is more demanding in L2 processing *per se,* see also Chapter 2.2.2.3. However, her data was on semantic processing, the current observation is made for syntactic processing.

may be seen as an instance of N400. Frisch & Schlesewsky (2001) report an N400 effect as a result of processing double nominative violations. However, they attribute the N400 effect to problems of the processing system to thematically hierarchize two identically case-marked NPs. They put forward that the N400 is enhanced only when the wrongly case-marked critical item NP is animate, while no enhancement is found for critical inanimate NPs. Thus, in the present study, any difficulties establishing a thematic hierarchy between both NPs, due to the animacy-marking of the second / critical NP, should be diminished. There are two possible interpretations for the occurrence of the N400 effect in native speakers' ERPs. First, it might be due to a priming effect, which is triggered by the stimulus material. Syntactically and semantically critical sentences were presented simultaneously. For both sorts of material, the second NP was the critical one, marking either a semantic or syntactic violation. Additionally, the participants were asked to judge each sentence for its semantic and syntactic structure. The task might have led participants to draw attention to the second NP, and consequently activation and retrieval of integration mechanisms were enhanced automatically. Second, this integration difficulty may be due to some mechanism of retrieval of the information that the first NP already bears nominative. According to prediction, this excludes any possibility of assigning nominative case to the critical item. Thus, it could be seen as an instance of N400, albeit it does not resemble its classic distribution in neither group. A closer look at the data of Frisch & Schlesewsky (2001) reveals a negativity effect for the double nominative violation that also appears globally distributed. The authors label this as N400, though they do not present confirmed statistical results on distributional features. Again, their globally distributed N400 is enhanced only when the critical item is animate, while no enhancement is found for the inanimate NPs. One reason for this difference might be attributable to the syntactic structures (i.e., word order) of the stimulus materials. With respect to the structure of the current study, the prediction that the first NP has to be assigned agent-role is much more enhanced, as for the material used by Frisch & Schlesewsky. They presented both NPs immediately following each other in a subordinate clause and prior to the verb. In addition to this enhanced structural complexity, the first NP was manipulated by altering nominative or accusative case-marking. The second NP was always marked nominative and served as a critical item. In contrast, the critical item of the present study is the NP that is altered by case-marking. Thus, the ERPs are measured directly on the violation, which might generally enhance the processing costs of integration, irrespective of the NP's animacy status. Albeit the lack of a clear explanation for the occurrence of the N400, it seems to reflect the activation of neural mechanisms that are different from those referred to the processing of semantic incongruities (see Chapter 4.2.3). Evidence for this claim is found in the ERP data elicited by L2 learners. While in Experiment 1 (Chapter 4.2.3) the negativity effect reflecting semantic processing costs occurred robustly for all L2 learners, the negativity effect in the present data is elicited only when AoA is rather early—i.e., growing AoA yields a gradual attenuation of the negativity. This result goes in line with the results of previous L2 ERP studies, where the N400 stayed absent in L2 learners as a possible influence of AoA. Again, this was the case for ill case-marked animate NPs (Mueller et al., 2007; Mueller, 2009). It was further suggested that increasing proficiency would yield a LAN referring to increased activation and retrieval of syntactic processing mechanisms (Mueller et al., 2005; Steinhauer et al., 2009). In the present data there is no indication at all for an upcoming LAN due to the improving proficiency. Osterhout et al. (2006) propose an N400 elicited by low-proficiency late L2 learners for syntactic subject-verb agreement violations indicating that generally low proficiency triggers N400 effects rather than P600 effects on syntactic violation (see also Steinhauer et al., 2009). Yet, as already indicated, the present data does not reveal any proficiency influence on the negativity at all. Against this background, then, the occurrence of the N400 in L2 learner's ERPs can only be explained tentatively. Furthermore, it can only be cautiously presumed why it is only determined by AoA. As already mentioned, due to the experimental design, semantic and syntactic violations are presented concurrently and the participants are asked to judge the sentences for their syntactic and semantic structure. Again, since both violations occurred at the same position in the critical sentences, integration mechanisms may have been evoked, in addition to the case-marking violation. That is, due to the task, enhanced integration mechanisms of the second NP might have been triggered and hence the present N400 effect reflects a strong correlation with the underlying demands as to the processing of semantic incongruity. Once more, there is a caveat to this account. If the occurrence of the N400 was confounded by semantic incongruity, the AoA influence is not explainable. The data of Experiment 1 suggests that integration mechanisms are robustly retrieved, irrespective of either AoA or proficiency. In the current data, there is a clear AoA influence in the fact that increasing AoA causes the occurrence of the N400 effect to gradually attenuate and eventually to disappear. This further suggests two things: First, the N400 represents the activation of neural mechanisms that gradually decline as an instance of AoA; second, neural mechanisms underlying the processing reflected in Experiment 1 are not the same underlying Experiment 2. I shall return to these two issues in the general discussion (Chapter. 5.2).

In summary and with respect to the expected ERPs elicited by processing the double nominative violations, it can be stated that L1 and L2 ERPs both reveal a biphasic N400-P600 pattern as soon as a case-violation occurs. The occurrence of this pattern in L2 ERPs is highly dependent on both AoA and proficiency. Yet, the characteristics of their influence diverge. The N400 effect is strongly AoA-influenced as the increasing AoA reduces its strength. The P600 seems to be more impacted by proficiency, in general. Yet, this positive proficiency influence is less exhaustive as AoA increases. Finally, both factors independently and interactively show a gradual influence. The current data does not suggest any threshold or discontinuity.

4.4 Experiment 3: NPI Licensing

As was already pointed out in Chapter 2.2.3, processing of German NPI 'jemals' ('ever') yields the activation of neural resources and hence the retrieval of semantic and syntactic processing mechanisms. Accordingly, there is consistent evidence for native speakers to elicit an N400-P600 processing pattern when the NPI is not appropriately licensed. Results of the two previous experiments (cf. Chapters 4.2.3 and 4.3.3) indicate that syntactic but not semantic L2 processing mechanisms may be affected by interacting AoA and proficiency influence. It shall be interesting to investigate whether this interaction also arises when the devotion of both semantic and syntactic processing mechanisms are required within one sentence. Therefore, the following experiment investigates ERPs as response to the processing of the German NPI 'jemals' ('ever') in licensed, compared with non-licensed, contexts by native German speakers and L2 learners with different AoA and proficiency levels.

4.4.1 Stimulus Materials and Hypotheses

In this third experiment the German NPI 'jemals' ('ever') is presented in either licensed or non-licensed contexts.[102] Sample sentences are displayed in Table 4-13 below.[103] For a full list of stimulus sentences, see appendix 5.3. Subject NP and direct object NP in the stimulus material both are masculine and animate. The critical item (NPI 'jemals' ('ever')) always directly follows the direct object NP. Hence, the linear structure of the sentence is subject NP (nominative, masculine, singular), auxiliary 'hat' ('has'), direct object NP (accusative, masculine, singular and definite), NPI and past participle of a transitive verb. The subject NP either includes a negative quantifier 'kein' ('no') or the definite article 'der' ('the'). Accordingly, only the subject NP including the negative

[102] In order to avoid prediction strategies analogous structures with the positive polarity item 'durchaus' ('definitely') were also presented. This should prevent the participants from predicting the occurrence of the NPI after reading the second NP (direct object).

[103] NPI stimulus material was partly adopted from Drenhaus et al. (2007) and modified to meet the criteria of stimulus material set out for the present study.

quantifier appropriately licenses the NPI. The definite article is considered the strongest licensing violation (for corresponding ERP results see Drenhaus et al., 2007; under review; Schuette, 2006). This might maximize the possibility that potential processing mechanisms of L2 learners are triggered, since in the mean they might not be familiar with these rather rare constructions relative to the L2 input (J. Blaszczak, personal communication July, 08. 2014).[104] Different from the stimulus material used in the former experiments, the underlying grammatical requirements for NPI licensing are quite distinct in German and Polish. Polish is a language with negative concord, whereas German is not. This means that in the appropriate structure of a Polish sentence like (5b), the verb as well as the direct object has to be indicated by a negative marker (J. Blaszczak, personal communication August, 20. 2014).

Table 4-13: Stimulus examples of licensed and non-licensed NPI conditions: NOM = nominative, ACC = accusative, '*'-indexed sentences entail NPI licensing failure.

Correct licensed NPI condition
Kein$_{NOM}$ Chemiker$_{NOM}$ hat den$_{ACC}$ Physiker$_{ACC}$ *jemals* geärgert.
No$_{NOM}$ chemist$_{NOM}$ has the$_{ACC}$ physician$_{ACC}$ *ever* annoyed.
Incorrect non-licensed NPI condition
*Der$_{\mathbf{NOM}}$ Chemiker$_{NOM}$ hat den$_{ACC}$ Physiker$_{ACC}$ *jemals* geärgert.
*The$_{NOM}$ chemist$_{NOM}$ has the$_{ACC}$ physician$_{ACC}$ *ever* annoyed.

(5)

a) Kein Lehrer hat den Schüler jemals bestraft
No teacher has the student ever punished

b) Żaden nauczyciel nie ukarał nigdy żadnego ucznia
No teacher not (has) punished never no student

[104] All L2 learners were familiar with the German word 'jemals' ('ever') according to the test of vocabulary knowledge (see Chapter 4.1.4).

Albeit the cross linguistic difference, on the basis of former results on the L1 and L2 processing of NPI structures (see Chapter 2.2.3) a biphasic N400-P600 pattern is expected to be revealed by native speakers' ERPs. That is, as soon as the processor encounters the non-licensed NPI, the N400 should be enhanced due to integration difficulties of the NPI into the prior sentence context. The subsequent P600 should reflect enhanced processing costs due to activation and retrieval of reanalysis / repair mechanisms. The latter are retrieved because of the structural unavailability of a proper licensor, which renders the structure ungrammatical. Expectations concerning ERP-responses to the processing of an unlicensed NPI elicited by L2 learners are difficult to predict on the basis of previous investigations. Therefore, predictions on whether AoA and / or proficiency show an influence on L2 learners ERPs can only be formulated cautiously. On the basis of the prior ERP results of the current study, the following predictions on influences on possible N400-P600 L2 processing patterns may be expected: Firstly, given the results of Experiment 1 (see Chapter 4.2.3), the expected N400 should not be much affected by neither factors since the activation of neural resources and the retrieval of mechanisms linked with semantic processing are challenged due to integration difficulties of 'jemals' ('ever'), for it is unexpected in a non-licensed context. Hence, the expected N400 should resemble brain responses associated with lexical rather than thematic L2 processing. Secondly, the P600 as an indicator of syntactic reanalysis / repair should be impacted by an interactive proficiency and AoA influence. Again, according to the results in Experiment 2 (see Chapter 4.3.3), the P600 elicited at early AoA should increase with improving proficiency. At later stages of AoA, this proficiency influence should reduce and hence the P600 should occur attenuated. This interactive influence is further expected to be continuous. There is one limitation to this hypothesis formulated for the expected P600. Hence, in Experiment 2, the occurrence of P600 infers to the clear ungrammaticality of the structure, whereas, in the present non-licensed NPI structure, the ungrammaticality is less overt in terms of morpho-syntactic violation. It, therefore, might also be the case that neural mechanisms concerning structural repair are less or even not at all activated by L2 learners. The results of behavioural and ERP data are presented in the following sections.

4.4.2 Behavioural Results: NPI Licensing

Mean accuracy ratings relative to NPI licensing conditions obtained for all participants are shown in Figure 4.20. Table 4-14 lists the average ratings for native speakers and L2 learners separately. Statistical analysis carried out for native speakers does not yield any differences across the accuracy ratings (model only had an intercept). The accuracy data of L2 learners show great variation for the ratings between both conditions. Statistically, the overall accuracy for the licensed condition is significantly higher than for the non-licensed condition ($F(1,3417) = 8.67$, $p = .003$)[105], indicating that L2 learners are less aware of the licensing failure. Proficiency ($F(1,3417) = 25.03$, $p < .001$)[106] and AoA ($F(1,3417) = 12.57$, $p < .001$)[107] both significantly influence the mean accuracy ratings. Additionally, their interaction is statistically reliable ($F(1,3417) = 14.26$, $p < .001$)[108], although it does not impact differences between the ratings across licensing conditions. Generally, at early AoA, the improving proficiency yields more accurate judgements. As AoA increases, this influence continuously reduces, which infers a loss of proficiency influence at latest AoA stages.

Table 4-14: Mean accuracy rates (in %) and RTs (in ms) revealed by native speakers and L2 learners for licensed and non-licensed NPI conditions with standard deviations in parentheses.

Group	Accuracy in %		Reaction Times in ms	
	licensed NPI	non-licensed NPI	licensed NPI	non-licensed NPI
Native Speakers (n = 18)	88.82 (31.53)	82.89 (37.68)	498.97 (294.25)	498.82 (270.49)
L2 learners (n = 60)	61.94 (48.56)	43.85 (49.63)	717.47 (491.9)	716.87 (477.48)

105 *Denominator lower bound* = 3207, $p = .003$.
106 *Denominator lower bound* = 3207, $p < .001$.
107 *Denominator lower bound* = 3207, $p < .001$.
108 *Denominator lower bound* = 3207, $p < .001$.

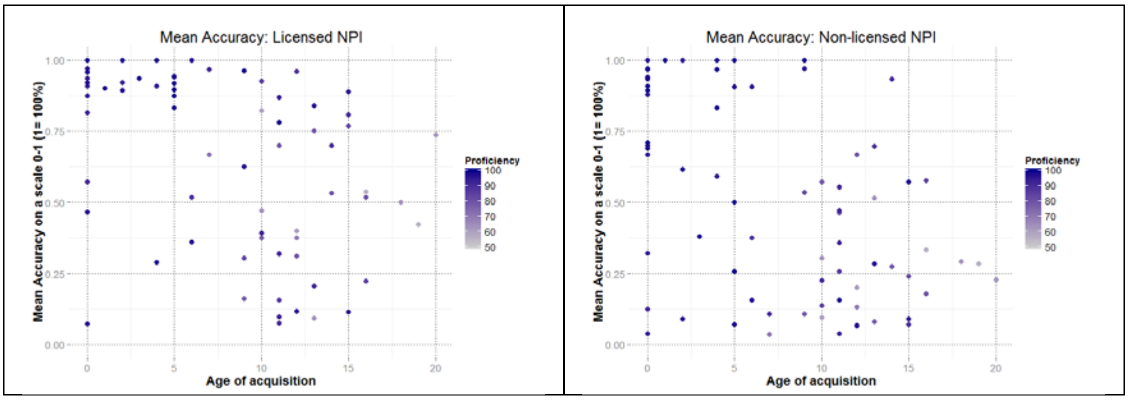


Figure 4.20: Mean accuracy rates (in %) for both, licensed (left) and non-licensed (right) NPI conditions for all participants (n = 78). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, accuracy = 100 % and shading = 90-100) are mapped only once.

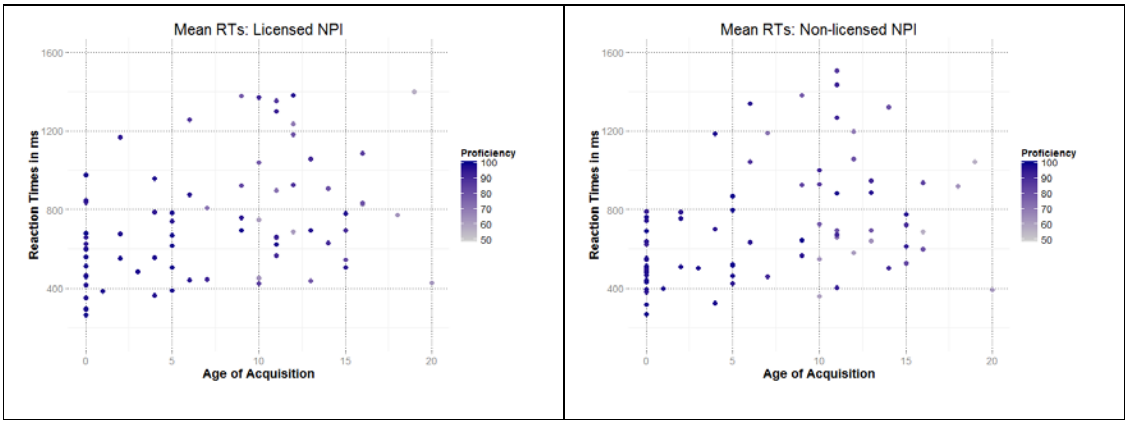


Figure 4.21: Mean RTs (in ms) for both, licensed (left) and non-licensed (right) NPI conditions obtained for all participants (n = 78). X-axis displays AoA in years. Shading of points indicates individual proficiency level, i.e. the measured C-Test score. Identical values (e.g., AoA = 0, RT = 600 ms and shading = 90-100) are mapped only once.

Average RTs for the NPI licensing conditions obtained for each participant are illustrated in Figure 4.21. Groups' averages are summarized by Table 4-14. Statistical results do not reveal any differences in the speed of responses for the licensed compared with the non-licensed condition in neither native speakers (model only had an intercept) nor L2 learners (Licensor as a fixed effect was removed from the model). With respect to the latter, improving proficiency generally increases mean RTs

($F(1,2682) = 6.93$, $p < .001$)[109], whereas increasing AoA gradually slows them down ($F(1,2682) = 12.24$, $p < .001$)[110]. However, neither of the factors determines differences between RTs across Licensor conditions. According to the great loss of data points due to the rather poor overall accuracy (see again Figure 4.19) a further analysis including the RTs of all judgements obtained for L2 learners was carried out. The outcomes show similar results relative to the influence of improving proficiency ($F(1,4428) = 4.97$, $p < .05$)[111] and growing AoA ($F(1,4428) = 6.1$, $p < .05$)[112]. Again, differences in the speed of responses across Licensor conditions are not reliable.

In summary, the behavioural results reveal that especially for L2 learners the mean accuracy is rather poor and more variant, and that RTs are generally slower when compared with those obtained for Experiments 1 (Chapter 4.2.2) and 2 (Chapter 4.3.2). This suggests that the L2 processing of NPI structures is more demanding than processing a double nominative violation or semantic incongruity in an L2. Further interpretations, e.g., on how these enhanced demands are influenced, will not be drawn given that current behavioural data are seen as an instance of offline measures.

4.4.3 ERPs: NPI Licensing

In total, 4,614 trials were computed for ERP averages. As many as 471 trials (10.21%) had to be excluded from averaging due to increased drift and blink artefacts. Albeit the poor accuracy (see Chapter 4.4.2) and hence missing succession in the desired sentence interpretation, the assumption still holds that the participants attentively processed the sentences. Therefore, ERPs of L2 learners were averaged, including all available segments. Again, for visual purpose only, the participants were grouped in accordance with their AoA and proficiency level (again see Chapter 4.2.3 for criteria set out for group separations). Figure 4.22 shows the difference waves in a time-window between

---

[109] *Denominator lower bound* = 2358, $p < .001$.
[110] *Denominator lower bound* = 2358, $p < .001$.
[111] *Denominator lower bound* = 4104, $p < .05$.
[112] *Denominator lower bound* = 4104, $p < .05$.

-100 ms and 1500 ms relative to the critical NPI 'jemals' ('ever'), which occurred at 0 ms.[113] Comparisons of ERPs across licensing conditions for each group are separately displayed in appendix 1.3.

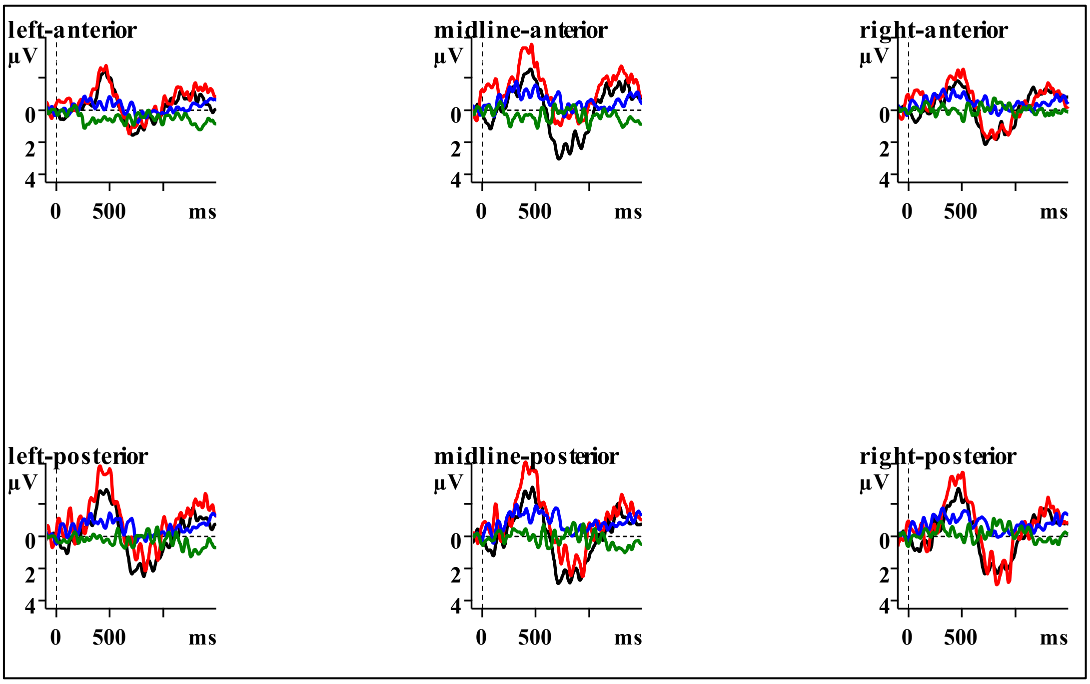


Figure 4.22: Difference wave forms of grand average ERPs time-locked to the critical NPIs according to the Licensor conditions by four groups separated only for visual purpose: black = Native Speakers, red = EAHP, blue = LAHP and green = LALP. Voltages are plotted on y-axis ranging from -5µV to +5µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. ROIs are labelled accordingly. Negative voltages are plotted up.

Native speakers show a large early negativity that appears the strongest between 400 and 500 ms, and is followed by a late positivity starting approx. around 700 ms. This processing pattern largely resembles the highly proficient L2 learners with early AoA. In comparison with the native speakers' ERPs, the early negativity appears enhanced, whereas the positivity appears reduced in strength and more right-laterally distributed. Regarding L2 learners with late AoA, ERPs do not show a late positivity. An early

[113] Grand average difference was calculated on the basis of the mean voltage of the licensed condition subtracted from the mean voltage of the non-licensed condition.

negativity is observable for those with late AoA and high proficiency only. However, it appears reduced in strength and slightly more right-laterally distributed. Low-proficiency L2 learners with late AoA do not show much difference between the ERPs of licensed vs. non-licensed 'jemals' ('ever'). With respect to visual inspection and in order to maintain comparability with the results of the two former experiments, two time-windows were cut and analysed: An early one between 400–500 ms and a late one between 800–900 ms. The consequent voltage difference maps of these two time-windows are displayed by Figure 4.23. The structure of the most complex statistical models corresponds to those used in Experiments 1 and 2 (see also Chapter 3.4.4 and appendix 4.3).

Native Speakers	EAHP	LAHP	LALP
AoA= 0	AoA < 8	AoA ≥ 8	AoA ≥ 8
proficiency > 80	proficiency > 80	proficiency > 80	proficiency ≤ 80

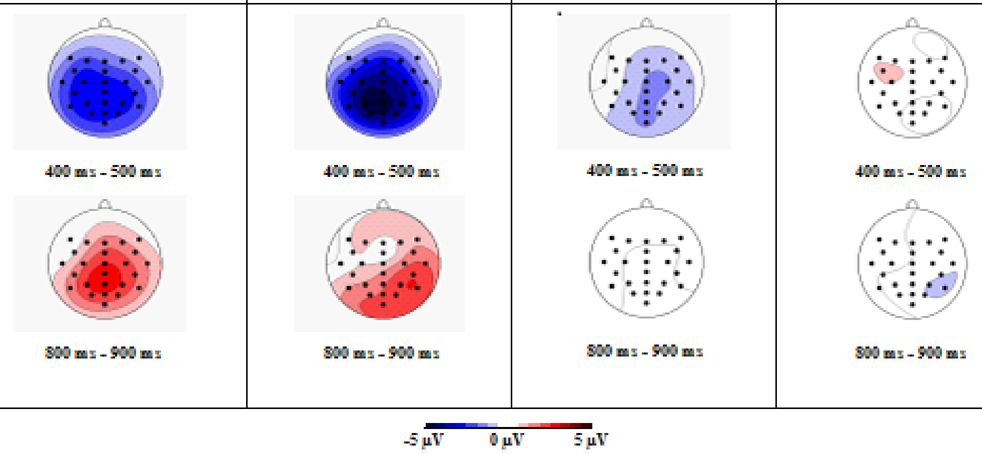


Figure 4.23: Difference voltage maps illustrating the mean differences of ERPs time-locked to the critical NPIs according to the Licensor conditions in two time-windows: 400–500 ms and 800–900 ms post stimulus. Differences are displayed for four groups separated only for visual purpose. Groups are indexed at the upper row of each column. Time range representing the average voltage difference for each head is labelled accordingly. Difference voltage range is plotted from -5µV (dark blue) to +5µV (dark red).

4.4.3.1 Early Time-Window: 400–500 ms

The analyses of ERPs by native speakers in the early time-window reveal a significant Licensor effect (average voltage difference = -2.32µV, $F(1,766) = 19.35$, $p < .001$[114]) and a significant effect for ROI ($F(5,766) = 7.21$, $p < .001$)[115]. The effect for Licensor is due to more negative mean potentials of the non-licensed condition. Furthermore, it occurs reliably stronger on posterior than anterior ROIs ($F(5,766) = 5$, $p < .001$)[116], indicating larger differences between licensor conditions on posterior ROIs. The results of the *post hoc* analyses of single ROIs are listed in Table 4-15 and reveal that the negativity effect is highly reliable on all scalp-sites and appears numerically more pronounced on posterior than anterior ROIs.

Table 4-15: ANOVA table of the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for native speakers. Diff = difference.

ROI	*Mean Diff in µV*	*Denominators upper-/lower-bound*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	-2.17	138/102	1	18.58	18.58	15.91	<.001	<.001
MIDLINE-ANTERIOR	-2-25	103/68	1	6.45	6.45	14.8	<.001	<.001
RIGHT-ANTERIOR	-1.6	138/102	1	7.95	7.95	14.12	<.001	<.001
LEFT-POSTERIOR	-2.73	138/102	1	24.92	24.92	20.78	<.001	<.001
MIDLINE-POSTERIOR	-2.79	103/67	1	8.95	8.95	14.06	<.001	<.001
RIGHT-POSTERIOR	-2.5	138/102	1	10.24	10.24	19.65	<.001	<.001

Statistical results of L2 learners' ERPs in the early time-window are listed in Table 4-16. The analysis reveals a significant effect for Licensor (mean voltage difference = -1.39µV), which is reliable on all ROIs. As can be seen in Table 4-17, presenting the results of planned *post hoc* ROI analyses the Licensor effect also is numerically larger on posterior than anterior electrodes.

---

[114] *Denominator lower bound* = 622, $p < .001$.
[115] *Denominator lower bound* = 622, $p < .001$.
[116] *Denominator lower bound* = 622, $p < .001$.

Table 4-16: ANOVA table of the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners according to the Licensor conditions (*denominator upper bound df* = 2550, *denominator lower bound df* = 2070).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
LICENSOR	1	31.55	31.55	31.45	< .001	< .001
ROI	5	128.03	25.61	25.53	< .001	< .001
PROFICIENCY	1	0	0	< 1		
AOA	1	0.11	0.11	< 1		
LICENSOR×ROI	5	44.21	8.84	8.81	< .001	< .001
LICENSOR×PROFICIENCY	1	22.31	22.31	22.24	< .001	< .001
LICENSOR×AOA	1	7.19	7.19	7.16	.01	.01
ROI×PROFICIENCY	5	10.31	2.06	2.06	.07	.07
ROI×AOA	5	8.35	1.67	1.66	.14	.14
PROFICIENCY×AOA	1	0.14	0.14	< 1		
LICENSOR×ROI×PROFICIENCY	5	14.21	2.84	2.83	.01	<.01
LICENSOR×ROI×AOA	5	23.84	4.77	4.75	< .001	< .001
LICENSOR×PROFICIENCY×AOA	1	8.96	8.96	8.93	< .001	< .001

Table 4-17: ANOVA table of the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners. Diff = difference.

ROI	*Denominators upper-/lower-bound*	*mean Diff in* μV	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	471/351	-0.78	1	7.46	7.46	6.32	.01	.01
MIDLINE-ANTERIOR	352/232	-1.53	1	9.4	9.4	18.63	<.001	<.001
RIGHT-ANTERIOR	462/342	-1.14	1	14.53	14.53	19.47	<.001	<.001
LEFT-POSTERIOR	464/344	-1.58	1	27.76	27.76	23.83	<.001	<.001
MIDLINE-POSTERIOR	340/220	-1.98	1	13.63	13.63	32.84	<.001	<.001
RIGHT-POSTERIOR	464/344	-1.64	1	23.58	23.58	31.23	<.001	<.001

With respect to the negativity's strength, both AoA and proficiency yield varying influence on its distribution and also show an interactive impact. The distributional variation due to proficiency influence on the negativity is illustrated in Figure 4.24. It shows that improving proficiency gradually enhances the strength of the negativity effect on all ROIs. The results of the *post hoc* analyses of the differences across licensor conditions due to proficiency influence for single ROIs are given in Table 4-18. As proficiency improves outcomes show significantly positive-going mean potentials for the licensed condition, whereas mean potentials for the non-licensed condition reliably are negative-going. This influence is highly significant for all scalp-sites and numerically largest on left-lateral ROIs.

Table 4-18: ANOVA table of proficiency influence on the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners.

ROI	*Denominators upper-/lower-bound*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	471/351	1	12.82	12.82	10.87	<.001	<.001
MIDLINE-ANTERIOR	352/232	1	7.11	7.11	14.09	<.001	<.001
RIGHT-ANTERIOR	462/342	1	8.92	8.92	11.95	<.001	<.001
LEFT-POSTERIOR	466/346	1	15.21	15.21	13.05	<.001	<.001
MIDLINE-POSTERIOR	343/223	1	6.3	6.3	14.6	<.001	<.001
RIGHT-POSTERIOR	466/346	1	10.92	10.92	14.46	<.001	<.001

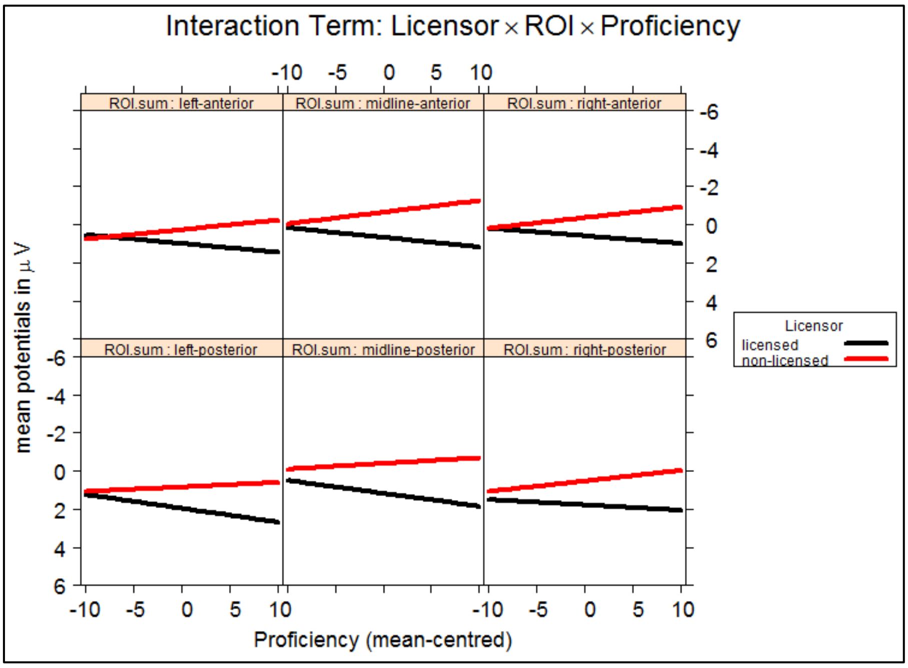


Figure 4.24: Plot of Interaction Term "Licensor × ROI × Proficiency" according to the time-window 400–500 ms. Proficiency data (mean-centred) is mapped on x-axis. Y-axis represents the mean voltages (range from 6µV to -6µV) of each condition: licensed NPI (black line) and non-licensed NPI (red line). Negative voltages are plotted up. Grids display ROIs (indicated as ROI.sum) and are labelled accordingly.

Figure 4.25 illustrates that increasing AoA also determines changes in the difference across the mean potentials of both conditions yielding an attenuation of the negativity's strength. This influence significantly varies with respect to distribution. Accordingly, separate analyses of each ROI are performed. Results are summarized by Table 4-19, and reveal negative-going mean potentials for the licensed condition and a positive shift of mean potentials for the non-licensed condition as AoA increases. Although this influence is reliable on all ROIs, it appears much stronger on posterior than anterior scalp-sites.

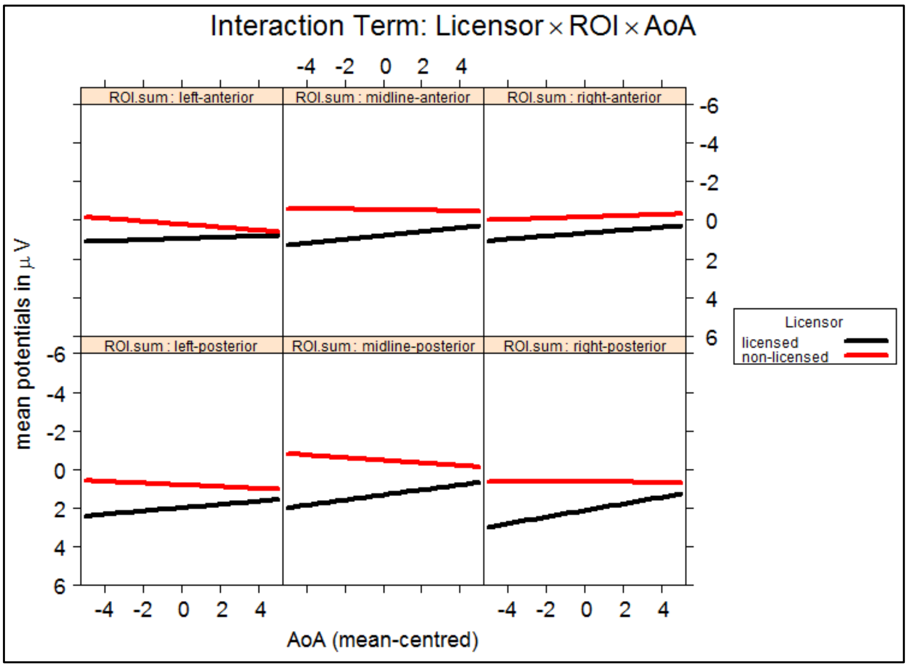


Figure 4.25: Plot of Interaction Term "Licensor × ROI × AoA" according to the time-window 400–500 ms. AoA data (mean-centred) is mapped on x-axis. Y-axis represents the mean voltages of each condition: licensed NPI (black line) and non-licensed NPI (red line). Negative voltages are plotted up. Grids display ROIs (indicated as ROI.sum) and are labelled accordingly.

Table 4-19: ANOVA table of AoA influence on the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for L2 learners.

ROI	*Denominators upper-/lower-bound*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	471/351	1	6.58	6.58	5.84	.02	.02
MIDLINE-ANTERIOR	352/232	1	6.36	6.36	12.6	<.001	<.001
RIGHT-ANTERIOR	463/343	1	5.01	5.01	6.75	.01	.01
LEFT-POSTERIOR	466/346	1	28.1	28.1	24.12	<.001	<.001
MIDLINE-POSTERIOR	343/223	1	11.37	11.37	26.36	<.001	<.001
RIGHT-POSTERIOR	466/346	1	17.83	17.83	23.61	<.001	<.001

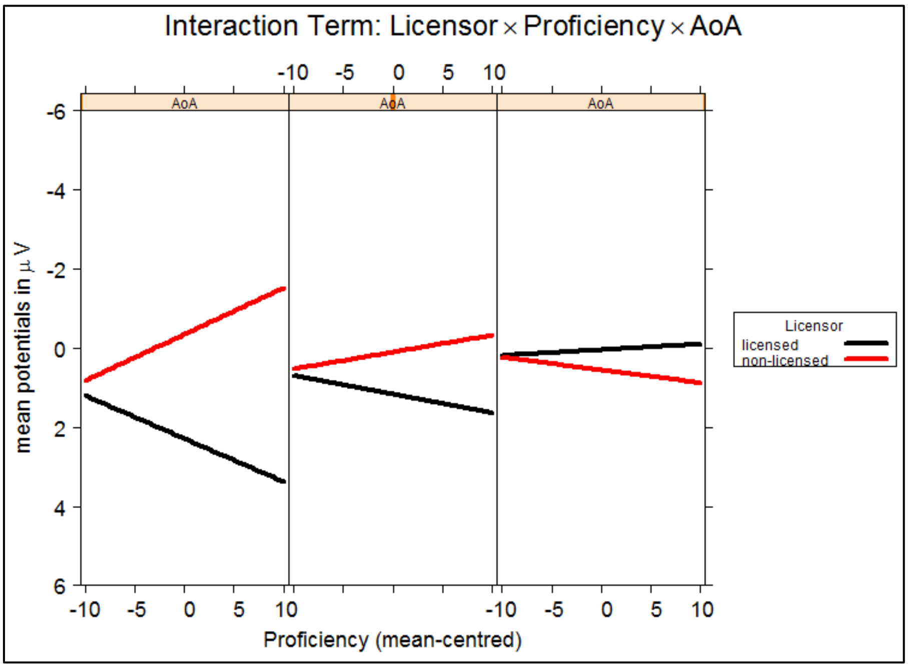


Figure 4.26: Plot of Interaction Term "Licensor × Proficiency × AoA" according to the time-window 400–500 ms. Proficiency data is mean-centred and mapped on x-axis. Y-axis represents the mean voltages of each condition: licensed NPI (black line) and non-licensed NPI (red line). Negative voltages are plotted up. Grids display three AoA stages (from mean-centred AoA data): left = earliest AoA, centre = middle AoA, and right = latest AoA.

The interactive influence of AoA and proficiency on the negativity is plotted in Figure 4.26. It indicates that for the non-licensed condition only at early AoA improving proficiency yields more negative-going mean potentials. With increasing AoA, this proficiency influence reduces gradually. At latest AoA stages, the mean amplitudes of the non-licensed condition are less affected by proficiency. With respect to the licensed condition, the mean potentials are more positive-going due to improving proficiency at earlier AoA. As AoA increases this influence deteriorates. At latest AoA stages, the proficiency influence does not appear to be dependable.

4.4.3.2 Late Time-Window: 800–900 ms

Statistical analysis of ERPs elicited by native speakers in the late time-window reveals significant main effects for Licensor ($F(1,765) = 14.59$, $p < .001$)[117], ROI ($F(5,765) = 16.12$, $p < .001$)[118] and their significant interaction ($F(5,765) = 9.23$, $p < .001$)[119]. The mean potentials for the non-licensed condition are more positive (1.16µV) than for the licensed condition (- 0.53µV). The outcomes of the *post hoc* analyses of single ROIs, given in Table 4-20, reveal that this positivity effect is consistent on all ROIs and numerically most pronounced on midline-posterior electrodes.

Table 4-20: ANOVA table of the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 800–900 ms conducted for native speakers. Diff = difference.

ROI	*denominators upper-/lower-bound*	*mean Diff in µV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	140/104	0.81	1	5.83	5.83	5.47	.02	.02
MIDLINE-ANTERIOR	106/70	2.03	1	11.4	11.4	9.71	<.001	<.001
RIGHT-ANTERIOR	140/104	1.39	1	8.48	8.48	9.13	<.001	<.001
LEFT-POSTERIOR	140/104	1.95	1	32.23	32.23	14.16	<.001	<.001
MIDLINE-POSTERIOR	105/69	2.68	1	16.54	16.54	21.86	<.001	<.001
RIGHT-POSTERIOR	140/104	1.98	1	26.06	26.06	15.25	<.001	<.001

117 *Denominator lower bound* = 621, $p < .001$.
118 *Denominator lower bound* = 621, $p < .001$.
119 *Denominator lower bound* = 621, $p < .001$.

Table 4-21 reports the statistical results relative to the ERPs elicited by L2 learners for the late time-window. The outcomes reveal a small mean difference of 0.33μV between licensor conditions—i.e., a positivity. However, statistically the positivity effect is not reliable. Distibutional differences in the strength of the positivity between ROIs do not yield any significant results either (see Table 4-22 for single ROI analyses that were performed due to planned comparisons). Yet, there is a significant interaction between Licensor and AoA indicating that with increasing AoA the positivity (effect) gradually attenuates. As is illustrated by Figure 4.27, this is due to significantly negative-going mean potentials of the non-licensed condition ($F(1,1278) = 4.91$, $p = .03$)[120], whereas mean amplitudes of the licensed condition are not reliably affected by AoA influence.

Table 4-21: ANOVA table of the average ERP amplitudes in the time-window 800–900 ms conducted for L2 learners according to the Licensor conditions (*denominator upper bound df* = 2551, *denominator lower bound df* = 2071)

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
LICENSOR	1	1.17	1.17	1.16	.28	.28
ROI	5	124.18	24.84	24.71	< .001	< .001
PROFICIENCY	1	0.14	0.14	< 1		
AOA	1	0	0	< 1		
LICENSOR×ROI	5	7.11	1.42	1.41	.22	.22
LICENSOR×PROFICIENCY	1	1.27	1.27	1.26	.26	.26
LICENSOR×AOA	1	8.93	8.93	8.89	< .001	< .001
ROI×PROFICIENCY	5	10.85	2.17	2.16	.06	.06
LICENSOR×ROI×PROFICIENCY	1	46.04	9.21	9.16	< .001	< .001

[120] *Denominator lower bound* = 1218, $p = .03$.

Table 4-22: ANOVA table of the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 800–900 ms conducted for L2 learners. Diff = difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/ lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	476/356	0.38	1	3.27	3.27	1.72	.18	.18
MIDLINE-ANTERIOR*	356/236	0.02	1	0	0	< 1		
RIGHT-ANTERIOR*	476/356	0.25	1	0.88	0.88	0.73	.39	.39
LEFT-POSTERIOR*	476/356	0.42	1	2.83	2.83	1.88	.17	.17
MIDLINE-POSTERIOR*	356/236	0.39	1	0.6	0.6	0.89	.34	.34
RIGHT-POSTERIOR*	476/356	0.5	1	1.52	1.52	1.5	.22	.22

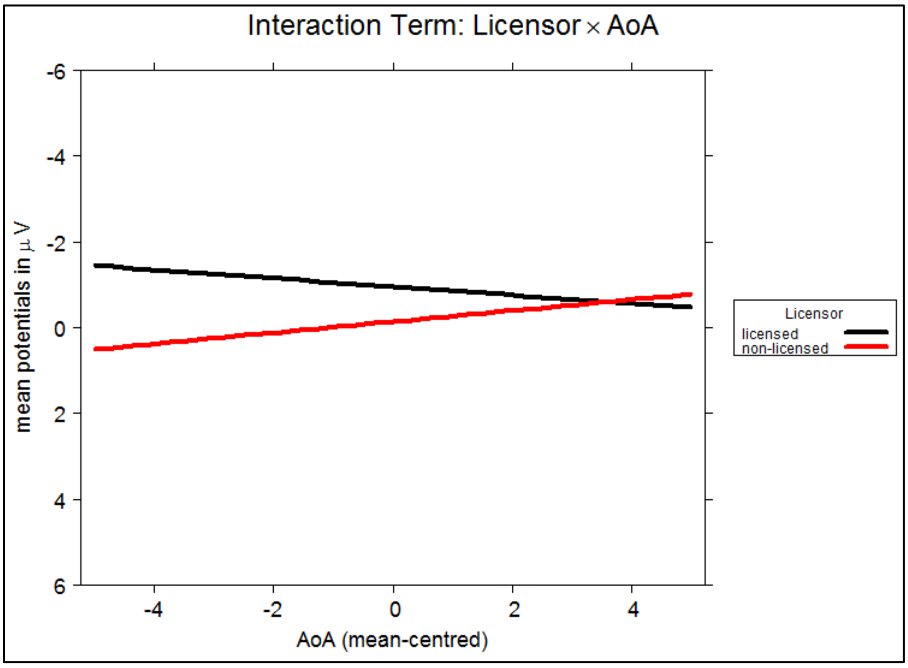


Figure 4.27: Plot of Interaction Term "Licensor × AoA" according to the time-window 800-1000 ms. AoA data is mean-centred and mapped on x-axis. Y-axis represents the mean voltages of each condition: licensed NPI (black line) and non-licensed NPI (red line). Negative voltages are plotted up.

Further, the three-way interaction between Licensor, ROI and proficiency, depicted in Figure 4.28 indicates that there are distributional differences relative to proficiency impact on the mean potentials across conditions yielding topographical variation in the strength of the positivity. Results of *post hoc* analyses of single ROIs are given in Table

4-23. They yield an enhancement of the positivity due to improving proficiency only on right-posterior scalp-sites. An interactive influence of both AoA and proficiency on the positivity elicited by L2 learners is not reported.

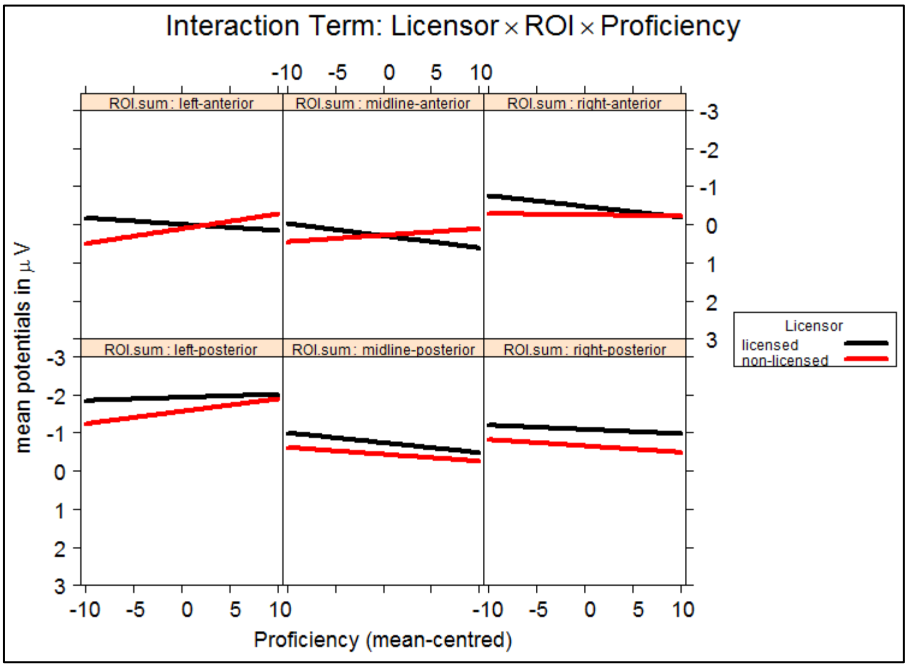


Figure 4.28: Plot of Interaction Term "Licensor × ROI × Proficiency" according to the time-window 800–900 ms. Proficiency data is mean-centred and mapped on x-axis. Y-axis represents the mean voltages of each condition: licensed NPI (black line) and non-licensed NPI (red line). Negative voltages are plotted up. Grids display ROIs (indicated as ROI.sum) and are labelled accordingly. Scaling of y-axis is reduced to increase readability.

Table 4-23: ANOVA table of proficiency influence on the Licensor effect in each ROI according to the average ERP amplitudes in the time-window 800–900 ms conducted for L2 learners. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*Denominators upper-/lower-bound*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	476/236	1	0	0	< 1		
MIDLINE-ANTERIOR*	347/111	1	0	0	< 1		
RIGHT-ANTERIOR	465/165	1	10.69	10.69	7.32	.01	.01
LEFT-POSTERIOR	466/166	1	6.92	6.92	4	.05	.05
MIDLINE-POSTERIOR	347/111	1	13.84	13.84	6.73	.01	.01
RIGHT-POSTERIOR	465/170	1	49.52	49.52	24.99	<.001	<.001

4.4.4 Summary and Discussion: NPI Licensing

In this third experiment, ERPs as response to L1 and L2 processing of German NPI 'jemals' within licensed and non-licensed contexts were investigated. Generally, the average ERPs of native speakers reveal a biphasic processing pattern when the NPI 'jemals' ('ever') is not appropriately licensed, namely an early negativity followed by a late positivity, as was expected. With respect to the stimulus material and the proposed hypotheses this result will be referred to as biphasic N400-P600 pattern. This finding corresponds to the results of former studies (see Chapter 2.2.3.1) and can be interpreted as follows: Whenever the language processor encounters the non-licensed NPI 'jemals' ('ever'), processing demands are enhanced due to difficulties of updating the context wherein the NPI cannot be integrated (N400 effect). Further, since the absence of an appropriate licensor renders such a sentence ungrammatical neural mechanisms linked with processes of structural repair / reanalysis are activated and retrieved (P600).

L2 learners' ERPs reveal differences compared to native speaker's ERPs. Although it is obvious that the underlying neural resources of both semantic and syntactic processing mechanisms are activated, the degrees of activation are constrained by both AoA and proficiency. With respect to the early negativity—i.e., N400 effect—increasing AoA generally yields its attenuation. This is due to changes of mean potentials of the licensed condition and the non-licensed condition. Especially on (right-) posterior scalp-sites, where the classic N400 effect is often found to be the largest, the influence of increasing AoA is clearly visible by the negative shift of mean potentials of the licensed condition, which indicates an increase of difficulty to generally integrate the NPI into the sentence context. Further, proficiency determines an enhancement of the N400 effect. This influence decreases as AoA increases. Generally, processing mechanisms of updating the context with new information (i.e. NPI) are activated and retrieved. AoA, then, is predictive for the activation enhancement due to NPI licensing failure. Improving proficiency further supports the activation enhancement, but only when AoA is rather early. Proficiency cannot compensate for AoA influence on later AoA stages. Consequently, the activation of the neural resources for syntactic mechanisms of repair reflecting the P600 is subject to AoA influence. That is, the P600 attenuates with increasing AoA. Moreover, improving proficiency is not really predictive for changes in the strength of the P600, albeit it claims an impact on its distributional differences as response to the processing of a non-licensed NPI. Accordingly, in order to reflect a biphasic N400-P600 ERP pattern, the neural resources of both semantic and syntactic processing mechanisms have to be activated. These activations are highly AoA-constrained in L2 processing. Besides, there seems to be a dependency of retrieval between semantic and syntactic processing mechanisms. Syntactic mechanisms will be retrieved only when the activation of neural resources related with semantic processing mechanisms is enhanced. In other words, the language processing system has to be sensitive to the NPI licensing-failure in order to show both brain responses as reflection of integration difficulty and subsequent brain responses associated with mechanisms of repair / reanalysis. If an L2 learner does not notice that an upcoming NPI is not appropriately licensed due to missing negation in the former sentence context and therefore does not reflect difficulties in updating the context with the NPI, there will not be any reason to further retrieve the processing mechanisms associated with repair or reanalysis of the sentence.

To sum up: According to the main objectives concerning the investigation of ERPs as response to L1 and L2 processing of NPI 'jemals' ('ever') in an non-licensed context (see Chapter 2.2.3), there are (i) differences between L1 and L2 processing patterns as response to the processing of the non-licensed NPI; (ii) there is no clear dominance for the activation of either semantic or syntactic L2 processing mechanisms. Yet, the activation of subsequent syntactic processing mechanisms requires the prior activation of semantic processing mechanisms; (iii) AoA is highly predictive in that enhanced processing costs due to integration difficulty and subsequent repair are reflected when AoA is rather early. The results do not suggest a threshold or discontinuity but a rather gradual attenuation of the corresponding ERP effects. With respect to the N400 effect, AoA is also predictive for proficiency impact—i.e., the earlier AoA, the more obvious the strengthening of the N400 effect due to the improving proficiency. This proficiency influence continuously reduces with increasing AoA. Finally, (iv) AoA and proficiency influences on the L2 processing of NPIs appear rather different from their influence elicited for isolated semantic and syntactic L2 processing. A detailed interpretation of this latter observation will be taken up in the general discussion below.

5 General Discussion

In the present thesis, ERPs as response to L1 and L2 processing of semantic incongruity, double nominative violation, and NPI licensing were investigated. These studies were intended to contribute to the observed inconsistency of AoA and proficiency influence as a source of differences between L1 and L2 processing mechanisms reflected by ERPs. Further, it was anticipated to gain more insights into the correlation between the factors AoA and proficiency (i.e., the older the AoA, the lower the proficiency) trying to unfold potential differences of their weighting with respect to their impact on L2 processing mechanisms.

In the following, the ERP results of native speakers and L2 learners separately will be summarized and further interpreted in the light of former findings.

5.1 L1 ERPs

The ERPs of native speakers largely reveal the processing patterns that were expected in accordance with the results of former studies for all three structures that were investigated. The processing of a semantically incongruent word enhances processing costs since the word cannot easily be integrated into former sentences' context. This integration difficulty is reflected by an enhanced N400 and refers to the greater activation of neural mechanisms that are required to update the previous context with the new word. Although the N400 effect elicited by native speakers in the present study on the processing of semantic incongruity appears rather small (as indicated by statistical analysis), it still is in accordance with the common findings relative to L1-semantic processing: It is enhanced when a word causes integration difficulty and it shows a right-posterior distribution, i.e. is largest on right-posterior scalp-sites.

Native speakers' ERPs as response to the processing of a double nominative violation yield a biphasic N400-P600 processing pattern. Although the N400 effect is rather unexpected, it may be explained in terms of activation and retrieval of neural mechanisms that are devoted to thematic processing. It means that due to the identical case-marking of both subject and object NP, the latter triggers enhanced activation of the neural resources associated with thematic processing mechanisms in order to resolve the case-marking failure by assigning thematic roles. This assignment, then, is necessary in order to clarify that subject and object NPs are in accordance with the former structural prediction due to the canonical word order and hence to structurally integrate the object NP into the prior sentence. However, this interpretation diverges from former interpretations of the thematic N400 effect as response to double nominative violations. Frisch & Schlesewsky (2001) argue that the N400 is enhanced only when animacy marking cannot aid to resolve the case-marking failure. They put forward, that if the animacy marking of identically case-marked NPs is different, it provides a hint to determine the thematic hierarchy of both NPs and hence resolve the case-marking failure. Further, if animacy marking of identically case-marked NPs also is identical, it creates an unresolvable conflict and only this unresolvable conflict enhances the N400. However, in the present study, the conflict concerning identical case-marking should be resolvable due to the different animacy marking of both NPs and hence the N400 should not have been enhanced in terms of Frisch & Schlesewsky's interpretation. Albeit the possibility of a task-related enhancement of the N400, as already indicated in Chapter 4.3.4, the enhanced N400 relative to the case-violation might have been triggered by the strong mismatch with the structural prediction. That is, the canonical word order assumes an upcoming NP that bears *no* nominative case-marking. When the upcoming NP is case-marked nominative the processing system has to activate additional thematic processing mechanisms in order to resolve the conflict caused by the mismatch with structural prediction. This additional activation then might be reflected by the enhanced N400. To explain the enhanced N400 (also) in terms of mismatch with structural prediction and hence, inferring a 'structural cloze probability' approach, expands the integrative account as well as the thematic account. That is, in order to elicit such an N400 effect, it is highly probable that additional or even different neural resources have to be activated and retrieved. Evidence for this claim is demonstrated by the corresponding L2 learners' ERP patterns. For the latter it seems that the devotion of these additional or even different neural resources is diverse with regard to the N400 effect elicited by the processing of semantic incongruity (see also Chapter 5.2 below). This structural prediction account has to be seen different from structural reanalysis (i.e., P600, see below). The former imply the activation of retrieval mechanisms due to the mismatch with previous prediction on the basis of nominal case marking (nominative marking vs. accusative marking). The latter reflect neural mechanisms of reanalysis that are activated in order to resolve and repair the case-marking failure.

Subsequent to the N400 effect, the ERPs of native speakers as response to the processing of double nominative violation reveal a P600, which largely resembles findings of former results with respect to strength and distribution. Therefore, the present interpretation of the P600 goes in line with previous ones in that it reflects the activation of neural resources retrieved for syntactic repair mechanisms, because the case-marking failure renders the structure ungrammatical.

A biphasic N400-P600 pattern is also revealed by native speakers' ERPs as response to the processing of an NPI that is not appropriately licensed. This finding resembles that of former results on the L1 processing of similar structures. Therefore, both brain responses are understood in the line with former interpretations. The N400 effect reflects enhanced processing costs due to the difficulty of integrating the NPI into the former context. Since there is no appropriate licensor, the upcoming NPI creates a strong mismatch as it is not predicted and, therefore, cannot easily be integrated into the sentence' context. The subsequent P600 reflects enhanced activation of neural resources linked with repair / reanalysis mechanisms. The structure renders ungrammatical because of the absent licensing element as the NPI is required to be in the scope of an appropriate licensor.

5.2 L2 ERPs

The ERPs of L2 learners as response to semantic incongruity processing reveal a robustly enhanced N400 that on average has a central distribution. This largely resembles the former findings on semantic L2 processing, indicating that L2 learners devote neural resources associated with the integration of a word into the former sentence context. In case of integration difficulty, the activation of resources relative to semantic processing mechanisms is enhanced. The N400 effect shows significant differences in scalp distribution in terms of anteriority and is numerically largest on midline-anterior scalp-sites. Lateral asymmetries are not found. In general, this confirms previous findings which have quite consistently reported a loss of hemispheric asymmetry and more left-lateralised N400 effects especially in comparison with L1 processing (e.g., Newman et al., 2012; Ojima et al., 2005; Weber-Fox & Neville, 1996). These comparative differences between the N400 distribution for L1 and L2 cannot be maintained by the present data. Moreno & Kutas (2005), also, do not report any lateral asymmetry in the N400 effect elicited by processing of neither the dominant nor the non-dominant language. The authors state that given their results and with respect to language no reliable comments can be made according to differences across functional brain organisation of L1 and L2. This, then, as well might be the case for the present data. In the broader time-window the present results reveal (only numerically!) a right-posterior scalp-distribution for native speakers' N400 effect, whereas for L2 learners the N400 effect is the greatest on midline-anterior scalp-regions. For the latter, again, this is similar to the findings of the study by Moreno & Kutas. The authors speculate that the processing of words for the L2 (by L2 learners) is more concrete than for the L1. This means that L2 learners are entangled to interpret a sentence context by rather concrete and imaginable than abstract features. Concrete words are more consistent and less language-specific in their use and, therefore, easier to interpret (cf. ibid. p.218). The stimulus material of the present study has not been controlled for concreteness of the critical words, yet it is conjecturable that the more frontally distributed N400 effect relates to more concrete processing of the nouns that were used as stimulus items (again see appendix 5.1 for the complete list of stimulus sentences).

One last important point concerning the L2 processing of semantic incongruity refers to the fact that neither AoA nor proficiency influence is observed. This is consistent with the general understanding that semantic processing mechanisms are intact and can fully be retrieved for the L2. However, this finding is different from the results of previous studies that relate AoA and / or proficiency to differences between ERPs elicited by L1 and L2 semantic processing (see Chapter 2.2.1). There are several reasons for the lack of any influence (of AoA and / or proficiency) in the current data: First, it might just be the case that there simply is no influence and that the neural resources considered to make sense of the sentence are available and can be fully retrieved (see above). This would also support the idea that shallow processing (see Chapter 2.3) is sufficient to semantically integrate a word into the former sentence' context and either AoA or proficiency play any important role. Second, it could very well be that the continuous treatment of both variables is responsible for the lack of statistical influence. Recall that most of the former studies have used a categorical design where L2 learners were separated into groups, and the variances and mean values of groups were compared. A closer look at the present data in Figure 4.9 and Figure 4.10 leaves the assumption that there may be some AoA and proficiency influence with regard to strength (AoA-related) and lateral asymmetry (proficiency-related) of the N400 effect. This, then, would be consistent with the results of Newman et al. (2012; again see Chapter 2.2.1.2) but cannot be reliably confirmed with respect to the present statistical design. Yet, present results reveal an AoA influence—though not on the N400 effect—on the mean potentials in general. Increasing AoA reveals a gradual positive shift of mean potentials. Hence, there is evidence for isolated influence of AoA that shows continuous character. Unfortunately, there is no clear explanation for this observation—i.e., positive-shift of mean potentials as a function of increasing AoA. Since there is no reliable explanation that may be related to the present finding, this question has to remain unsolved.

ERPs as response to L2 processing of double nominative violations reveal a biphasic N400-P600 pattern which is similar to that found for native speakers (see Chapter 5.1 above for interpretations of the two brain responses). However, the biphasic pattern is highly dependent on both the AoA and proficiency level of the L2 learner. With respect to the strength of the ERP effects elicited by L2 learners, results show a gradual decrease of the N400 effect as a function of AoA. Interestingly, this is due to a continuous negative shift of the ERP response to the case-congruent condition. Similar results have been reported by Hahne (2001) and also Mueller (2009). The authors suggest that demands to integrate any word into the former context are enhanced during L2 processing which, hence, yields an attenuation of the N400 effect. In these studies, again, groups were compared categorically and the group of L2 learners who all had a late AoA did not show any N400 effect (relative to the N400 effect elicited by the ERPs of native speakers). The present finding is consistent with the results of these studies and may be expanded to the fact that the attenuation of the N400 effect due to the more negative-going mean potentials of the case-congruent condition is determined by AoA and appears continuously. However, this suggests a somewhat different interpretation. It was indicated above that the N400 as response to case-violation elicited by L1 speakers is seen as a reflection of enhanced activation of additional processing mechanisms caused by the mismatch with structural prediction (case-violation on the second NP). Further, the enhanced N400 cannot be explained in terms of either thematic conflict resolution or mere integration difficulties. As to the latter, AoA influence on the N400 effect due to the case-violation should have resembled the N400 effect elicited due to semantic incongruity (Experiment 1). Since this is not the case, there is the strong indication that the neural resources for semantic and structural integration mechanisms are different. Moreover, the ability to retrieve semantic resources is less (or even not) AoA influenced whereas the ability to retrieve resources associated with structural integration gradually declines as AoA increases. In other words, it seems that enhanced L2 processing demands due to conflicting structural prediction become less available with increasing AoA. Conceivably, with increasing AoA L2 learners do not build a structural prediction as to case-marking. In the light of shallow processing (as described in Chapter 2.3) the current finding puts forward that the neural mechanisms that have to be activated and retrieved in order to integrate a word semantically are different from those that relate to structural integration. That is, shallow processing seems to be appropriate enough to process a semantic incongruity in a native like manner, whereas shallow processing is not sufficient in terms of native like processing and constrained by AoA when the processor comes across a double nominative violation. This indicates that L2 learners with late AoA process such structures on the mere basis of lexical-semantic information. It may be argued, then, that for L2 processing the determiner information (case-marking) is not as much predicted as it is for native processing. L2 processing relies on the lexical information in order to integrate the noun into the former context. In terms of shallow L2 processing this observation might infer L1 influence. Since Polish does not have any determiner system L2 learners might not be aware of the importance of the German determiner as a case marker. However, this would additionally suggest that improving proficiency should have impacted the N400 and that L2 processing may be become more native-like due to improving grammar development. In contrary, results suggest a strong AoA influence. Feasibly, the AoA impact indicates that the according neural resources cannot be activated and retrieved in L2 processing because of maturational restrictions. Interestingly, in case of late controlled processing mechanisms of syntactic reanalysis/repair —as revealed by the P600— L2 learners are able to activate and retrieve those. This further indicates that improving proficiency may compensate for AoA impact concerning the activation and retrieval of the according processing mechanisms (see below) and, that mechanisms of structural integration and repair are of different neural resources.

For L2 learners, the subsequent P600 as reflection to retrieve mechanisms associated with structural repair / reanalysis processes is less sensitive to AoA influence. The P600 appears at all stages of AoA. However, there is also proficiency influence suggesting that increasing proficiency enhances the P600 at late AoA. This goes in line with former findings on L2 processing of case-violations and also with the idea that AoA-related effects on syntactic L2 processing (i.e. absence of P600) may be compensated by improving proficiency (Steinhauer et al., 2009; see also Chapter 2.2.2.3). That is, the neural resources related to mechanisms of repairing an ill-formed structure are activated as soon as the L2 learner is aware of the fact that there is a syntactic violation in the input, irrespective of AoA. This finding supports the shallow structure hypothesis in that the processing mechanisms associated with full parsing strategies may become more automatized in L2 as proficiency and, consequently, L2 grammar improves. However, the present data suggest that AoA influence cannot be fully compensated as the decreasing exhaustiveness of proficiency influence correlates with increasing AoA of the L2 learner. I.e. at later stages of AoA improving proficiency is less efficient concerning the ability to activate and retrieve the neural resources of syntactic repair mechanisms. Hence, structural processing continuously becomes more demanding and shallower for L2 processing with increasing AoA and gradually limits the compensatory influence of proficiency.

The L2 processing of an NPI like German 'jemals' ('ever') in a non-licensed context also reveals a biphasic N400-P600 ERP pattern as is found in native speakers' brain responses. Likewise to AoA and proficiency influence on the L2 processing of case-violations, a native-like N400-P600 pattern only occurs in brain responses of L2 learners with early AoA and who are highly proficient. However, the single and interactive influence of both factors is different compared with that on case-violation (see above) indicating either the devotion of different neural resources, or different degrees of activation, or both. The present results indicate that the neural resources responsible for the integration of a non-licensed NPI into the former context can be activated and retrieved irrespective of AoA. Yet, at later AoA, L2 leaners need to have obtained a rather high level of L2 proficiency in order to activate these integration mechanisms. That is, the L2 learner has to be aware of the fact that the NPI is not licensed and therefore, not acceptable in a non-licensed sentence context. If the L2 learner does not notice the licensing failure, the according neural mechanisms are not activated and, hence, not retrieved. Further, proficiency influence to compensate AoA effects (i.e. absence of N400 effect) is limited with respect to the activation of integration mechanisms. It can be fully exhausted only at earlier AoA. Again, this interactive impact is gradual and does not yield discontinuity. In order to robustly reflect enhanced processing mechanisms associated with syntactic repair (i.e. P600), AoA influence seems to be rather strong, although improving proficiency, too, yields enhanced activations of syntactic repair mechanisms. There is no interactive influence as could be demonstrated for the repair mechanisms due to case-violations (see above). One reason for this might be that on surface, the structural ungrammaticality of the NPI licensing failure is less salient than for case-violation. In order to activate neural mechanisms of syntactic repair, the L2 learner does not only have to notice that the NPI is not acceptable within a sentence lacking a prior licensing element (e.g. negation). The L2 learner also has to be aware that the missing licensor renders the structure ungrammatical. This ability to devote the neural resources and processing mechanisms seems to gradually attenuate as AoA increases. Hence, the enhanced processing mechanisms are devoted to the reflection of integration difficulty to a greater extent than to the need of syntactic repair. This further suggests that in order to retrieve the processing mechanisms of a syntactic repair / reanalysis, L2 learners heavily rely on the degree of salience as to the (syntactic) violation. In terms of shallow processing the present findings on L2 NPI processing suggest that licensing conditions are less shallow. A non-licensed NPI is not understood as grammatical violation in the same way as a case marking violation.

5.3 Conclusion

The present thesis has presented ERP research on the differences between the processing of German as a native language and as L2. It has investigated how AoA and / or proficiency impact(s) on the processing of L2 and how such influences can be accounted for the potential differences between L1 and L2 processing patterns. Different from a bulk of former research on the issue, this paper has investigated the possible interactive influence of AoA and proficiency on (semantic and syntactic) L2 processing and also if these potential influences appear gradually. In general, the results imply that the neural dissociation between semantic and syntactic processing mechanisms is active in L2 learners. Furthermore, demands as to the processing system—i.e., which neural mechanisms need to be activated—differentiate as they heavily depend on the specific structure and / or the violation that has to be processed. ERP responses of native speakers that throughout different violations (semantic, syntactic and semantic-syntactic) superficially resemble in terms of strength and distribution (i.e. N400 effect) still reflect retrieval from different neural resources that are activated. This confirms the idea of neural dissociation for language processing. Further evidence is found in ERPs elicited by L2 learners in such a way that their brain responses are differently influenced by AoA and / or proficiency yielding different degrees of limitation to activate and hence retrieve the consequent processing mechanisms. With respect to the weighting of both influencing factors, the results indicate that AoA more than proficiency triggers differences between L1 and L2 processing patterns as well as between L2 processing patterns. Further, although improving proficiency may compensate for AoA effects, the degree of the compensation itself is limited by increasing AoA. Finally, the present results strongly suggest that changes in the ability to devote neural resources to L2 processing due to AoA and / or proficiency develop gradually. In the present data there is no indication of any discontinuous thresholds or borders set by AoA or proficiency level that would abruptly allow or deny—e.g., brain responses being similar to those elicited by native processing. This supports and extends recent suggestions of stage-like syntactic development in an L2 (e.g., McLaughlin et al., 2010; Osterhout et al., 2006; Steinhauer et al., 2009; Tanner et al., 2013) and endorses the concept of continuous influence of AoA and proficiency (e.g., Newman et al., 2012). Yet, with respect to former findings, the present data cannot account for any evidence on whether AoA or proficiency determines differences in the semantic processing of L2. As already indicated, it seems that the present results better confirm former evidence on syntactic processing, including the possibility to extend former interpretations. There might be one reason for this asymmetry in the present results and their relation to previous outcomes. Previous studies on semantic processing have reported either AoA or proficiency influence. Many studies have not strictly controlled for the collinearity of both factors and their results might therefore be subject to the omitted variable bias. In the present study, both factors were included in the analysis and the results show that there is no clear evidence for either factor to influence L2 semantic processing. This, then, indirectly supports the diverse outcomes of prior studies and puts forward that the results, also, largely depend on the statistical and factorial design in question. Prior studies on syntactic double nominative violations have more strictly controlled for the collinearity reducing the potential omitted variable bias. Accordingly, as already discussed, the present results (on syntactic L2 processing) largely confirm former outcomes. Thus, this further supports the statistical design used in the present study since present outcomes underline former results due to more controlled statistics and further extend the assumptions of an interactive influence of both factors by accounting for their collinearity. However, there are some limitations of the present results and their interpretation (see above), and more inquiry on this issue is needed.

Finally, I would like to touch on four points that might be interesting for further research: First, in the present study L1 influence has not been issued. Within L2 processing research there is an ongoing critical discussion on whether the (structures of a) native language may have impact on L2 processing (see e.g., review by van Hell & Tokowitz, 2010 and Papadopoulou & Clahsen, 2003 for diverging arguments supporting L1 influence). Conceivably, with respect to NPI-processing, Polish and German structures and processing demands largely diverge (see Chapter 4.4.1). It might be interesting to see whether the ERPs of L2 learners of German with an L1 that has similar semantic and syntactic NPI restrictions (e.g., English 'No man has ever disturbed the teacher.') show less differences compared with the ERPs of native speakers of German. Although highly speculative, it is reasonable to expect that the exhaustiveness of L2 proficiency may be less limited as a function of the structural L1-L2 similarity.

Second, the present results of Experiments 2 and 3 show somewhat reverse results concerning the characteristics of AoA and proficiency influence on N400 and P600 components elicited by L2 processing. Taking into consideration that the processing demands reflected by the N400 are different between experiment 2 and 3 might be a piece to the puzzle why AoA and proficiency reveal different influence characteristics. Still, several questions concerning the neural basis of the N400 remain open (see below). However, the processes of the activation and retrieval of neural resources reflecting the P600 were assumed to resemble for both the processing demands in Experiments 2 and 3. Whereas there is interactive influence on the P600 when a structural violation is apparent (Experiment 2), it is not significant when the violation is not as outward (Experiment 3). Questions that arise out of this observation may be: Do these differences indicate varying neural resources or degrees of activation and retrieval concerning the P600? How does this result contribute to the recent discussion of whether and to what degree the P600 can be considered a reflection of mere syntactic processing mechanisms (see Footnote 22)?

Third, the current results indicate that the neural resources underlying the N400 and / or the mechanisms to activate and retrieve them depend on the structure that is processed. Albeit the integration view holds for all of the presented processing patterns, distributional and strength differences (for L1 and L2 processing), as well as varying AoA and proficiency influence (for L2 processing) put forward the need to learn more about the nature of the N400. Is the strength of the N400 additive? I.e. does the processing of a structural violation reveal a stronger N400 effect than a semantic incongruity? What exactly does the varying influence of AoA and proficiency point to? Do L2 learners in contrast to native speakers devote different and / or subserving processing resources?

And fourth, characteristics of AoA and proficiency influence on electrophysiological evidence as presented in this thesis may contribute to the development and modification of neurocognitive models of L2 processing as e.g. the shallow structure hypothesis. Insights might be gained with respect to the determination of what makes a structure shallower than another one? Further, issues may be raised such as whether there is a mapping of a structural implication onto processing mechanism(s)? Are there processing mechanisms that remain shallow for L2 processing whereas others do not?

Resümee

Prozesse des Sprachverstehens unterliegen der sog. Dissoziation zwischen semantischer und syntaktischer Verarbeitung. Dies wurde u.a. mittels der Methode der ereigniskorrelierten Potenziale (EKP) anhand des sog. Verletzungsparadigmas untersucht. Dabei werden die EKP-Muster für die Verarbeitung einer korrekten/wohlgeformten Struktur mit denen einer verletzten/anomalen Struktur verglichen. Es konnte gezeigt werden, dass deutsche Muttersprachler/innen beim Lesen semantisch anomaler Sätze wie z.B. „Der Autor schreibt einen Stuhl an seinen Freund" neuronale Prozesse der Sprachverarbeitung aktivieren, die sich von denen beim Lesen syntaktisch anomaler Sätze wie z.B. „Der Autor schreiben einen Brief an seinen Freund" unterscheiden. Bei Ersteren zeigt sich im EKP eine verstärkte N400-Komponente, die allgemein mit erhöhtem Aufwand der semantischen Verarbeitung assoziiert wird (Kutas & Hillyard, 1980). Der erhöhte Verarbeitungsaufwand aufgrund einer syntaktischen Anomalie wird im EKP durch eine Verstärkung der P600-Komponente reflektiert (Gouvea et al., 2012; vgl. Friederici, 2002).

Vor diesem Hintergrund stellt sich die Forschung die Frage ob die Verarbeitung einer Zweitsprache (L2) auch dieser Dissoziation unterliegt. Es gilt zu untersuchen, ob bei der L2-Verarbeitung auch auf unterschiedliche neuronale Ressourcen zugegriffen wird, in Abhängigkeit davon, ob es sich um eine semantische oder syntaktische Anomalie handelt. Eine Möglichkeit zur Beantwortung dieser Frage bietet die Untersuchung von Gemeinsamkeiten und Unterschieden in der Verarbeitung (z.B. mittels EKP-Verarbeitungsmuster) zwischen Muttersprache (L1) und L2. Bisherige EKP-Ergebnisse drängen zu der Annahme, dass auch die Verarbeitung einer L2 der o.g. Dissoziation unterliegt, es jedoch sowohl quantitative als auch qualitative Unterschiede zwischen L1- und L2-Verarbeitungsmustern gibt (z.B. Hahne, 2001; Moreno et al., 2008). Ferner zeigt sich vor allem in den Unterschieden zwischen L1- und L2-Verarbeitungsmustern, dass eben diese Verarbeitungsunterschiede je nach Struktur (semantisch oder syntaktisch) von verschiedenen Einflüssen abhängig sind und dies zudem unterschiedlich stark. Diese Einflüsse sind naturgemäß für die L1-Verarbeitung nicht beobachtbar.

Zwei Faktoren, die die Verarbeitung einer L2 beeinflussen und somit auch zu Unterschieden verglichen mit L1-Verarbeitungsmustern führen, sind das Erwerbsalter und der Kenntnisstand. Der Einfluss des Erwerbsalters geht auf die Annahme zurück, dass es ein bestimmtes biologisches Zeitfenster gibt, in dem eine (zweite) Sprache gelernt/erworben werden muss. Findet der Erwerb innerhalb dieses Zeitfensters statt, so ist die sog. muttersprachliche Kompetenz erreichbar. Beginnt der Erwerb außerhalb dieses Zeitfensters ist muttersprachliche Kompetenz nicht mehr erreichbar (vgl. Lenneberg, 1967; Penfield & Roberts, 1959). Letzteres sollte dann anhand von Unterschieden z.B. zwischen L1- und L2-Verarbeitungsmustern zu beobachten sein. Der zweite Faktor, der die L2-Verarbeitung beeinflusst, ist der jeweilige Kenntnisstand in der L2. Es wird davon ausgegangen, dass ein niedriger Kenntnisstand in der L2 zu größeren Unterschieden zwischen L1- und L2-Verarbeitungsmustern führt, als ein hoher Kenntnisstand (z.B. Newman et al., 2012). Ein für die Untersuchung dieser Einflüsse auf die Verarbeitung problematisches Phänomen ist, dass beide Faktoren, Erwerbsalter und Kenntnisstand, korrelieren. Experimentelle Befunde haben gezeigt, dass in Abhängigkeit vom steigenden Erwerbsalter der Kenntnisstand in der L2 sinkt (Flege et al., 1999; Johnson & Newport, 1989). Problematisch daran ist, dass z.B. bei Untersuchungen von Verarbeitungsmustern von L2-Lernern mit spätem Erwerbsalter nicht genau vorhergesagt werden kann, ob wirklich das Erwerbsalter für potenzielle Unterschiede zwischen L1- und L2-Verarbeitung verantwortlich ist, oder nicht doch / auch der Kenntnisstand. Um dieses Problem zu vermeiden oder zumindest zu verringern, muss einer der beiden Faktoren (statistisch) kontrolliert sein. Solche Faktorenkontrolle wird bevorzugt durch Kategorisierung dieser beiden Variablen realisiert, die eigentlich charakteristisch kontinuierlich sind. Man gruppiert z.B. L2-Lerner/innen mit frühem und spätem Erwerbsalter und setzt relativ arbiträre Grenzen für z.B. früh und spät. Zuzüglich kontrolliert man den Kenntnisstand, der sich zwischen den beiden Gruppen nicht unterscheiden sollte. So umgeht man das Problem der

Korrelation, verliert jedoch auch sehr viel (statistisches) Potenzial, das sich aus der Korrelation für die Untersuchung ergeben könnte (vgl. z.B. Clarke, 2005).

<u>Fragestellung:</u> Die vorliegende Dissertation beschäftigt sich mit der Problematik dieser Korrelation. Neben den Fragen, welchen Einfluss das Erwerbsalter und der Kenntnisstand auf die Verarbeitung der L2 nehmen und welche daraus resultierenden Unterschiede sich bezüglich der Dissoziation zwischen der Verarbeitung semantischer und syntaktischer Strukturen ergeben, widmet sich diese Arbeit zentral der Frage, ob die Einflüsse von Erwerbsalter und Kenntnisstand unabhängig oder interaktiv auftreten und ob sie eher auf einen kontinuierlichen Verlauf oder sichtbare Grenzen hinweisen.

Um mehr Einblick in diese Problematik zu bekommen, thematisiert die vorliegende Dissertation Gemeinsamkeiten und Unterschiede in der Verarbeitung zwischen L1 und L2 und ferner den Einfluss von Erwerbsalter und Kenntnisstand auf eben diese anhand dreier EKP-Experimente zur Verarbeitung des Deutschen als L1 und L2 (die kritischen Elemente sind jeweils kursiv, siehe Beispiele (1-3) unten): Das erste Experiment untersucht die Verarbeitung von semantisch inkongruenten Wörtern im Satzkontext, siehe unten Beispiel (1). In einem zweiten Experiment wird die Verarbeitung einer syntaktischen Kasusverletzung (Doppel-Nominativ-Verletzung) wie in (2) untersucht (siehe unten). Im dritten Experiment wird die Verarbeitung einer Struktur untersucht, die sowohl semantische als auch syntaktische Ansprüche an den Verarbeitungsapparat stellt, die Lizenzierung des negativ polaren Elements (im Weiteren NPI) „jemals" wie in (3), (siehe auch z.B. Drenhaus et al., 2005). Die Untersuchung der Verarbeitung dieser dritten Struktur wurde gewählt, um zu sehen, ob sich die potenziellen Einflüsse, die sich isoliert für die Verarbeitung von Strukturen in (1) und (2) zeigen, schemenhaft auch auf kombinierte Ansprüche an den Verarbeitungsapparat herausstellen.

(1) Der Mann schreibt *den Roman/*den Stuhl* und erhält einen Preis

(2) Der Mann schreibt *den_{ACC} $Roman_{ACC}$ / *der_{NOM} $Roman_{NOM}$* und erhält einen Preis

(3) Kein/*Der Autor hat den Roman *jemals* geschrieben

Wie bereits oben erwähnt, ist nach derzeitigem Forschungstand davon auszugehen, dass die Verarbeitungsprozesse einer L2, wie die der L1, sensitiv gegenüber der Dissoziation zwischen semantischer und syntaktischer Verarbeitung sind. Weiterhin konnte wiederholt gezeigt werden, dass die Unterschiede zwischen der semantischen L1- und L2-Verarbeitung nicht so stark sind, wie die zwischen der syntaktischen L1- und L2-Verarbeitung. Befunde belegen, dass L2-EKPs für die Verarbeitung semantischer Anomalien robust eine erhöhte N400-Komponente reflektieren und dass die erhöhten N400-Komponenten sich relativ gering von denen unterscheiden, die in L1-EKPs reflektiert sind (z.B. Ardal et al., 1990; Newman et al., 2012; Ojima et al., 2005; Weber-Fox & Neville, 1996). Die angesprochenen geringen Unterschiede der N400-Komponenten zwischen L1 und L2 sind meist quantitativ und präsentieren sich anhand unterschiedlicher Distribution, Stärke oder Latenz. Die Verteilung der N400-Komponente evoziert in L2-EKPs ist meist bilateral (z.B. Weber-Fox & Neville, 1996). Des Weiteren wurden für L2-EKPs N400-Komponenten berichtet, die niedrigere Amplituden und längere Latenzen (z.B. Ardal et al., 1990; Newman et al., 2012) zeigen. Jedoch bleibt aufgrund bisheriger Ergebnisse unklar, ob diese Unterschiede auf das Erwerbsalter (Weber-Fox & Neville, 1996) und/oder den Kenntnisstand (Ardal et al., 1990, Newman et al., 2012; Ojima et al., 2005) zurückzuführen sind.

L2-EKP-Verarbeitungsmuster evoziert durch syntaktische Anomalien weisen stärkere Unterschiede verglichen mit L1-EKP-Verarbeitungsmustern auf. Anders als zwischen semantischer L1- und L2-Verarbeitung findet man hier sowohl quantitative als auch qualitative Unterschiede. So zeigen L2-EKPs im Gegensatz zu L1-EKPs meist eine stark reduzierte P600-Komponente (vgl. Steinhauer et al., 2009). Distributive und zeitliche Unterschiede werden weniger beleuchtet. Frühe EKP-Komponenten die mit erhöhten syntaktischen Verarbeitungsansprüchen assoziiert werden (z.B. LAN, vgl. Friederici, 2002) werden in L2-EKPs meist nicht reflektiert, was als qualitativer Verarbeitungsunterschied supponiert wird. Als Grund für diese sowohl quantitativen als auch qualitativen Unterschiede in der syntaktischen Verarbeitung wird der Einfluss des Erwerbsalters angenommen. Das bedeutet zum einen, dass Aktivierung und Zugriff auf neuronale Ressourcen gemäß der syntaktischen Verarbeitung durch spätes Erwerbsalter eingeschränkt operieren. Zum anderen scheint der Einfluss des Erwerbsalters größere Unterschiede in der syntaktischen als in der semantischen L2-Verarbeitung hervorzurufen. Jedoch konnte anhand longitudinaler Untersuchungen (z.B. McLaughlin et al., 2010) und Studien mit Miniatursprachen (z.B. Mueller et al., 2007) gezeigt werden, dass ein steigender Kenntnisstand die angesprochenen Erwerbsalterseffekte kompensieren kann. Dies deutet auf eine Interaktion beider Faktoren hin (vgl. Steinhauer et al., 2009).

Veröffentlichte Untersuchungen zur L2-Verarbeitung mit noch komplexeren Verarbeitungsansprüchen, wie NPI-Lizenzierung (siehe (3)), sind mir bisher nicht bekannt.

Allgemein lässt sich also feststellen, dass die Aktivierung und der Zugriff auf neuronale Ressourcen bezüglich semantischer Verarbeitungsprozesse robuster und weniger anfällig bezüglich der Einflussfaktoren sind als für syntaktische Verarbeitungsprozesse. Unklar bleibt, welcher der beiden Faktoren Einfluss nimmt und wie dieser gewichtet ist. Für die semantische Verarbeitung gibt es Befunde sowohl für einen Erwerbsalters- als auch Kenntnisstandeinflusses. Bezüglich der syntaktischen Verarbeitung zeigt sich ein interaktiver Einfluss beider Faktoren. Zudem bleibt das Problem der Korrelation, das in bisherigen Studien mehrheitlich dahingehend kontrolliert wurde, dass der Einfluss einer der beiden Faktoren (statistisch) ausgeschaltet wurde und dementsprechend ein Bias ob der Untersuchungsfrage entsteht (entweder nach dem Einfluss des Erwerbsalters *oder* des Kenntnisstands). In der vorliegenden Dissertation wird versucht, diesen Bias zu vermeiden. Beide Einflussfaktoren und somit auch ihre Korrelation gehen als kontinuierliche Variable in die Untersuchung mit ein. Mithilfe neuerer statistischer Methoden, die gemischten Modelle (z.B. Baayen et al.,2008) und mit der Aufhebung der Kategorisierung beider Faktoren ist es ein Ziel dieser Arbeit zu zeigen, ob sich unterschiedliche Gewichtungen der beiden Faktoren in ihrem Einfluss auf die L2-Verarbeitung zeigen und weiterhin ob sich dieser auch unterschiedlich für semantische, syntaktische und kombinierte Verarbeitungsmuster zeigt.

EKP-Daten von 20 deutschen Muttersprachler/innen und 60 L2-Lerner/innen des Deutschen mit Polnisch als L1 wurden erhoben. Das Erwerbsalter der L2-Lerner/innen reicht von 0-20 Jahre. Der Kenntnisstand (erhoben mittels eines standardisierten C-Tests (Klein-Braley & Raatz, 1982) rangiert in einem Bereich zwischen 53% und 99% gemäß der Einteilung des Kenntnisstands nach dem gemeinsamen europäischen Referenzrahmen für Sprache (vgl. Council of Europe, 2011). Alle deutschen Muttersprachler/innen berichteten als Erwerbsalter 0 Jahre und erreichten im C-Test > 93%. Zwei Korrelationstests (gerechnet mit und ohne Werte der Muttersprachler/innen) zeigen jeweils eine hoch signifikante Korrelation beider Einflussfaktoren.

<u>Experiment 1:</u> Die Analyse der EKP-Daten zeigt einen negativeren Verlauf der inkongruenten Bedingung im Zeitfenster zwischen 400 und 500 ms sowohl für Muttersprachler/innen als auch für L2-Lerner/innen. Dies wird als eine Verstärkung der N400-Komponente gedeutet. Der N400-Effekt ist statistisch in der Gruppe der Muttersprachler/innen nur auf rechts-parietalen Elektroden signifikant. Diese für den N400-Effekt typische Distribution unterstützt trotz des Fehlens des globalen Haupteffekts die Interpretation, dass die Verarbeitung eines semantisch inkongruenten Wortes im Satz im Vergleich zur Verarbeitung eines kongruenten Wortes erhöhte Ansprüche stellt, die durch die verstärkte N400-Komponente reflektiert werden. Die Analyse der L2-EKP-Daten im gleichen Zeitfenster demonstriert einen robusten N400-Effekt, der auf zentral-parietalen Elektroden numerisch am größten ist. Laterale Unterschiede in der Distribution zeigen sich nicht. Dieses Ergebnis unterstreicht bisherige Befunde, dass während der L2-Verarbeitung neuronale Ressourcen, die verantwortlich für die Integration eines Wortes in den bisherigen Satzkontext sind, robust aktiviert werden können. Im Falle einer Integrationsschwierigkeit wird diese Aktivierung erhöht. Die Analyse weist weder auf einen Einfluss des Erwerbsalters noch des Kenntnisstands hin. Diese Beobachtung kann leider nicht zur Beantwortung der Frage beitragen, welcher Einflussfaktor mehr Gewicht in der semantischen L2-Verarbeitung zeigt; sie geht gleichwohl konform mit dem allgemeinen Verständnis, dass

semantische Verarbeitungsmechanismen in der L2-Vearbeitung intakt sind und wirksam operieren.

<u>Experiment 2:</u> L1-EKPs bezüglich der Verarbeitung von Doppel-Nominativ-Verletzungen reflektieren ein biphasisches N400-P600 Verarbeitungsmuster. Im Zeitfenster zwischen 400 und 500 ms zeigt sich ein negativerer Kurvenverlauf wenn die Kasusmarkierung verletzt ist. Dieses Ergebnis ist unerwartet, kann aber gemäß erhöhter Ansprüche der thematischen Verarbeitung erklärt werden (z.B. Frisch & Schlesewsky, 2001). Durch identische Kasusmarkierung an Subjekt und Objekt (beide Nominativ) sind die Prozesse der syntaktischen (Subjekt, Objekt) und thematischen Rollenzuweisung (Agens, Patiens) erschwert. Dadurch ist eine Verstärkung dieser Verarbeitungsprozesse notwendig, um zu garantieren, dass die Besetzung von Subjekt- und Objektposition mit der strukturellen Vorhersage einhergehen und das Objekt (mit Patiensrolle) auch als solches – trotz Nominativmarkierung – in die Struktur integriert werden kann. Dem N400-Effekt folgend weisen die EKPs eine späte Positivierung mit zentral-parietaler Verteilung im Zeitfenster zwischen 800–900 ms auf. Diese Positivierung wird als P600-Effekt gedeutet, der als Konsequenz erhöhter Verarbeitungsmechanismen auftritt, die aktiviert werden müssen, um die fehlerhafte (ungrammatische) Struktur zu reparieren. Die L2-EKP Daten zeigen auch ein biphasisches N400-P600 Verarbeitungsmuster. Dieses Muster ist ähnlich dem der L1-EKPs. Jedoch zeigen sich sowohl für die frühe Negativierung als auch für die späte Positivierung Einflüsse von Erwerbsalter und Kenntnisstand. Die Analyse des frühen Zeitfensters (400–500 ms) zeigt keinen signifikanten N400-Effekt. Jedoch gibt es distributive Unterschiede in der Stärke: der N400-Effekt ist numerisch auf den Mittellinien ROIs am stärksten. Weiterhin ergibt die Analyse eine Interaktion der Stärke des N400-Effekts mit Erwerbsalter: mit steigendem Erwerbsalter wird der N400-Effekt kontinuierlich schwächer. Dies lässt vermuten, dass Verarbeitungsansprüche, die aufgrund des Konflikts der strukturellen Vorhersage und dem Auftreten einer Kasusverletzung erhöht sind, mit ansteigendem Erwerbsalter graduell weniger kompensiert werden können. Ein Einfluss des Kenntnisstands wurde nicht festgestellt.

Im späten Zeitfenster (800–900 ms) zeigt sich ein signifikanter Haupteffekt, der auf eine erhöhte Positivierung für die Kasusverletzungsbedingung zurückgeht und als P600-Effekt interpretiert wird. Der P600-Effekt präsentiert sich weniger anfällig gegenüber dem Einfluss des Erwerbsalters als der N400-Effekt. Weiterhin demonstrieren die Daten, dass ein steigender Kenntnisstand kontinuierlich den P600-Effekt vergrößert. Interessanterweise ist die Stärke dieses Einflusses (Kenntnisstand) abhängig vom Erwerbsalter. Diese Ergebnisse werden dahingehend interpretiert, dass Aktivierung und Zugriff neuronaler Ressourcen für Prozesse bezüglich der Reparatur einer kasusverletzten Struktur mit steigendem Kenntnisstand besser operieren. Dieser Kenntnisstandeinfluss nimmt mit steigendem Erwerbsalter ab. Mit anderen Worten, die strukturelle Verarbeitung einer Kasusverletzung wird mit zunehmendem Erwerbsalter anspruchsvoller und der kompensatorische Einfluss des Kenntnisstands ist vom Erwerbsalter beschränkt.

<u>Experiment 3:</u> Die EKP Daten der Muttersprachler zeigen ein N400-P600 biphasisches Verarbeitungsmuster wenn das NPI nicht adäquat lizensiert ist. Dies geht einher mit früheren Ergebnissen (z.B. Drenhaus et al., 2005) und legt nahe, dass es zu erhöhten Integrationsschwierigkeiten des NPI in einen nicht-lizensierten Kontext kommt, die durch den N400-Effekt reflektiert werden. Die Reanalyse der daraus resultierenden Ungrammatikalität des Satzes spiegelt die erhöhte Positivierung (P600-Effekt) wider. Ein biphasisches N400-P600 Verarbeitungsmuster präsentiert sich auch in den EKPs der L2-Lerner. Jedoch sind sowohl Stärke als auch Distribution beider Effekte abhängig von Erwerbsalter und Kenntnisstand. Im frühen Zeitfenster (400–500 ms) zeigt sich eine Verstärkung des N400-Effekts als Funktion des steigenden Kenntnisstands. Dieser Einfluss nimmt jedoch mit zunehmendem Erwerbsalter ab. Mechanismen bezüglich der Aktivierung und des Zugriffs auf neuronale Ressourcen die aufgrund der Integration eines NPIs in einen nicht-lizenzierten Kontext erhöht sind, können demnach mit ansteigendem Erwerbsalter weniger gut vom höheren Kenntnisstand kompensiert werden. Zusammenhängend damit zeigt sich im späten Zeitfenster (800–900 ms) eine kontinuierliche Abschwächung der P600 als Funktion steigenden Erwerbsalters. Dies

deutet auf einen Zusammenhang zwischen dem Auftreten beider Effekte hin, in dem die L2-Verarbeitung nur dann eine P600 reflektiert (Aktivierung und Zugriff auf neuronale Ressourcen bezüglich Reparaturmechanismen), wenn vorausgehend auch eine N400 robust evoziert wurde (Aktivierung und Zugriff auf neuronale Ressourcen hinsichtlich Integrationsprozesse).

Fazit: Die Ergebnisse der drei EKP-Experimente zeigen, dass die L2-Verarbeitung der Dissoziation zwischen semantischer und syntaktischer Verarbeitungsprozesse generell unterliegt. Wie stark diese Dissoziation aktiviert wird bzw. werden kann ist anhängig von beiden Faktoren, Erwerbsalter und Kenntnisstand. Obwohl beide Faktoren korrelieren, können gemäß dem gewählten Design solide Aussagen über die Gewichtung beider Faktoren gemacht werden. Es scheint, als ob das Erwerbsalter allgemein den stärkeren Einfluss auf die Aktivierung und den Zugriff neuronaler Ressourcen ausübt. Zeigt sich einen Einfluss des Kenntnisstands, so ist dessen Stärke/Gewichtung durch steigendes Erwerbsalter limitiert. Des Weiteren ist eine Symmetrie im Hinblick auf die Dissoziation zwischen semantischer und syntaktischer Verarbeitung zu beobachten. Bei syntaktischen Verletzungen werden thematische Verarbeitungsressourcen nur dann aktiviert, wenn die für die Prozesse von Reparatur/Reanalyse verantwortlichen Verarbeitungsmechanismen robust zugreifbar sind. Bei NPI Konstruktionen, kann auf syntaktische Verarbeitungsressourcen bezüglich Reparatur/Reanalyseprozesse nur dann zugegriffen werden, wenn auch solche verantwortlich für Integrationsmechanismen zugreifbar sind. Alle Einflüsse zeigen sich kontinuierlich. Die statistischen Ergebnisse liefern diesbezüglich keinen Hinweis auf kategoriale Abgrenzungen – weder für den Einfluss des Erwerbsalters (z.B. früh vs. spät) noch für den des Kenntnisstands (z.B. hoch vs. tief).

References

Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception and linguistic scrutiny. *Language Learning*, *59(2)*, 249-306.

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csáki (Eds.), *2nd International Symposium on Information Theory, Tsahkadsor, Armenia, USSR, September 2-8, 1971* (pp. 267-281). Budapest: Akadémiai Kiadó.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* 19 (6), 716-723.

Ardal, S., Donald, M.W., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain & Language*, *39(2)*, 187-205.

Baayen, R. H. (2010). A real experiment is a factorial experiment? *The Mental Lexicon*, *5*, 149-157.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Cognitive Neuroscience*, *59(4)*, 390-412.

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3(2)*, 12-28.

Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology*, *37(1)*, 13-20.

Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: An ERP comparison. *Journal of Cognitive Neuroscience*, *17(1)*, 137-53.

Bates, D. M., Meachler, M., Bolker, B., & Walker, S. (2014). *lme4*: *Linear mixed-effects models using Eigen and S4*. R package version 1.1-7. URL: http://CRAN.R-project.org/package=lme4.

van Berkum, J. J., van den Brink, D., Tesink, C. M., Kos, M., & Hagoort, P. (2008). The neural integration of speaker and message. *Journal of Cognitive Neuroscience*, *20(4)*, 580-591.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161-181). Mahwah, NJ: Lawrence Erlbaum Associates.

Birdsong, D. P. (1999). Introduction: Whys and why nots of the critical period hypothesis for second language acquisition. In D. P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 1-22). Mahwah, NJ: Lawrence Erlbaum Associates.

Birdsong, D. P. (2006). Age and second language acquisition and processing. A selective overview. *Language Learning*, *56 (s1)*, 9-49.

Birdsong, D. P. (2014). The critical period hypothesis for second language acquisition: Tailoring the coat of many colors. In M. Pawlak & L. Aronin (Eds.), *Essential topics in applied linguistics and multilingualism: Studies in honor of David Singleton* (pp. 43-50). Heidelberg: Springer.

Birdsong, D. P., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, *44(2)*, 235-249.

Bott, O. (2009). *The processing of events* (revised version of doctoral dissertation). John Benjamins (*Linguistics Today 162*), Amsterdam.

Brouwer, H., Fitz, H., & Hoeks, J. C. (2012). Getting real about semantic illusions: Rethinking the functional role of the P600 in language comprehension. *Brain Research*, *1446*, 127-143.

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33(2)*, 261-304.

Christian, W. (1984). Das Elektroenzephalogram (EEG) im höheren Lebensalter. *Nervenarzt*, *55*, 517-524.

Clahsen, H. & Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics, 27(1)*, 3-42.

Clahsen, H. & Felser, C. (2006b). Continuity and shallow structures in language processing. *Applied Psycholinguistics*, *27(2)*, 107-126.

Clarke, K. A. (2005). The phantom menace: Omitted variable bias in econometric research. *Conflict Management and Peace Science*, *22*, 341-352.

Coulson, S., King, J. W., & Kutas, M. (1998). Expect the unexpected: Event-related brain response to morphosyntactic violations. *Language and Cognitive Processes*, *13(1)*, 21-58.

Council of Europe (2011). *Common European Framework of Reference for: Learning, Teaching, Assessment.* Council of Europe.

Dambacher, M., Kliegl, R., Hofmann, M., & Jacobs, A. (2006). Frequency and predictability effects on event-related potentials during reading. *Brain Research*, *1084,* 89-103.

Deacon, D., Grose-Fifer, J., Yang, C. M., Stanick, V., Hewitt, S., & Dynowska, A. (2004). Evidence for a new conceptualization of semantic representation in the left and right cerebral hemispheres. *Cortex*, *40(3)*, 467-78.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22(4)*, 493-533.

Diedenhofen, B. (2013). *cocor: Comparing correlations*. R package version 1.0-0. URL: http://r.birkdiedenhofen.de/pckg/cocor/.

Domke, J. (2012). Eine EKP-Studie zum kindlichen Zweitspracherwerb. In K. Spalek, K. & J. Domke (Eds.), *Sprachliche Variationen, Varietäten und Kontexte: Festschrift für Rainer Dietrich* (pp. 187-202). Tübingen: Stauffenburg.

Drenhaus, H., Blaszczak, J., & Domke, J. (under review). The "fine structure" of polarity licensing from a psycholinguistic perspective: An ERP study on NPI-processing in German.

Drenhaus, H., Blaszczak, J., & Schuette, J. (2007). Some psycholinguistic comments on NPI licensing. In E. Puig-Waldmüller (Ed.), *Proceedings of Sinn und Bedeutung 11*, *Barcelona*, *Universitat Pompeu Fabra*, 180-193.

Drenhaus, H., & beim Graben, P. (2012). Ereigniskorrelierte Potenziale (EKPs). *Zeitschrift für Germanistische Sprachwissenschaft*, *40(1)*, 1-30.

Drenhaus, H., beim Graben, P., Saddy, D. & Frisch, S. (2006): Diagnosis and repair of negative polarity constructions in the light of symbolic resonance analysis. *Brain and Language*, *96(3)*, 255-268.

Drenhaus, H., Saddy, D. J., & Frisch, S. (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser & M. Reis (Eds.), *Linguistic Evidence: Empirical, theoretical, and computational perspectives* (pp. 145-165). Berlin: Mouton de Gruyter.

Engel, U., Rytel-Kuc, D., Cirko, L., & Debski, A. (1999). *Deutsch-polnische kontrastive Grammatik (Bd. 1 & 2)*. Heidelberg: Julius Groos.

Eubank L., & Gregg, K. R. (1999). Critical periods and (second) language acquisition: Divide et impera. In D. P. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 65-99). Mahwah, NJ: Lawrence Erlbaum Associates.

Fagan, S. M. B. (2009). *German: A linguistic introduction*. Cambridge: Cambridge UP.

Federmeier, K. D. (2007). Thinking ahead: The role and roots of prediction in language comprehension. *Psychophysiology*, *44(4)*, 491-505.

Flege, J., Yeni-Komshian, G., & Liu, S. (1999). Age constraints on second language learning. *Journal of Memory and Language*, *41(1)*, 78-104.

Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, *8(15)*, 1-27. URL: http://www.jstatsoft.org/v08/i15/.

Fox, J., & Hong, J. (2009). Effect displays in R for multinomial and proportional-odds logit models: extensions to the effects package. *Journal of Statistical Software*, *32(1),* 1-24. URL: http://www.jstatsoft.org/v32/i01/.

Frenzel, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Conflicts in language processing: A new perspective on the N400-P600 distinction. *Neuropsychologia*, *49(3)*, 574-579.

Friederici, A. D. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50(3)*, 259-281.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6(2)*, 78-84.

Frisch, S. (2000). Verb-Argument-Struktur, Kasus und thematische Interpretation beim Sprachverstehen (revised version of doctoral dissertation). MPI Series in Cognitive Neuroscience, 12. Dresden.

Frisch, S., & Schlesewsky, M. (2001). The N400 reflects problems of thematic hierarchizing. *Neuroreport*, *12(15)*, 3391-3394.

Frisch, S., & Schlesewsky, M. (2005). The resolution of case conflicts: A neurophysiological perspective. *Cognitive Brain Research*, *25(2)*, 484-498.

Garnsey, S. M. (1993). Event-related brain potentials in the study of language: An introduction. *Language and Cognitive Processes*, *8(4)*, 337-356.

Gelman, A. & Hill, J. (2006). *Data analysis using regression and multilevel / hierarchical models*. Cambridge: Cambridge UP.

Gouvea, A., Phillips, C., Kazanina, N., & Poeppel, D. (2010). The linguistic processes underlying the P600. *Language and Cognitive Processes*, *25(2)*, 149-188.

beim Graben, P., & Kurths, J. (2003). Detecting subthreshold events in noisy data by symbolic dynamics. *Physical Review Letters*, *90*, 100602.

Grotjahn, R. (2014), The C-Test bibliography: version January 2014. In R. Grotjahn (Ed). *Der C-Test: Aktuelle Tendenzen/The C-Test: Current trends* (pp. 323-365). Frankfurt am Main: Lang.

Grotjahn, R., Klein-Braley, C., & Raatz, U. (2002) C-Test: an overview. In J. A. Coleman, R. Grotjahn & U. Raatz (Eds.), *University language testing and the C-Test* (pp. 93-114). Bochum: AKS-Verlag.

Gunter, T. C., Friederici A. D., & Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *Journal of Cognitive Neuroscience*, *12(4)*, 556-568.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15(6)*, 883-899.

Hagoort, P., Baggio, G., & Willems, R. M. (2009). Semantic unification. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 819-836). Cambridge, MA: MIT Press.

Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, *8(4)*, 439-483.

Hagoort, P., Hald, L. A., Bastiaansen, M. C. M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, *304*, 438-441.

Hahne, A. (2000). Differences in semantic and syntactic processing in foreign language acquisition. *Journal of Cognitive Neuroscience*, (Suppl.), 51.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, *30(3)*, 251-266.

Hahne, A., & Friederici, A. D. (1999). Electrophysiological evidence for two steps in syntactic analysis: Early automatic and late controlled processes. *Journal of Cognitive Neuroscience*, *11(2)*, 194-205.

Hahne, A., & Friederici, A. D. (2001). Processing a second language: Late learners' comprehension strategies as revealed by event-related brain potentials. *Bilingualism: Language and Cognition*, *4*, 123-141.

Hahne, A., & Friederici, A. D. (2002). Differential task effects on semantic and syntactic processes as revealed by ERPs. *Cognitive Brain Research*, *13(3)*, 339-356.

Handy, T. C., Nagamatsu, L., Mickleborough, M., & Liu-Ambrose, T. Y. L. (2009). Statistical strategies for translational ERP studies. In T. C. Handy (Ed.). *Brain signal analysis: Advances in bioelectric and biomagnetic methods* (pp. 225-247). Cambridge, MA: MIT Press.

Harrell, F. E. Jr., with contributors from Charles Dupont and many others (2014). *Hmisc: Harrell Miscellaneous*. 3.14-5. URL: http://CRAN.R-project.org/package=Hmisc.

van Hell, J. G., & Tanner D. (2012). Second language proficiency and cross-language lexical activation. *Language Learning*, *62(Suppl.2)*, 148-171.

van Hell, J. G., & Tokowitz, N. (2010). Event-related brain potentials and second language learning: Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, *26(1)*, 43-74.

van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18(7)*, 1181-1197.

van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research*, *22(2)*, 241-255.

Hilton, H. (2014). Oral fluency and spoken proficiency: Considerations for research and testing. In P. Leclerq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 proficiency: Perspectives from SLA* (pp. 27-53). Bristol: Multilingual Matters.

Hyltenstam, K., & Abrahamsson, N. (2003). Maturational constraints in SLA. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 539-588). Malden, MA: Blackwell.

Jäger, F. (2008). Categorical data analysis: away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Cognitive Neuroscience*, *59(4)*, 434-446.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21(1)*, 60-99.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87(4)*, 329-354.

Kaan, E. (2007). Event-related potentials and language processing. A brief introduction. *Language and Linguistics Compass*, *1(6)*, 571-591.

Kaan, E., Harris, A., Gibson, E., & Holcomb, P. (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes*, *15(2)*, 159-201.

Klein-Braley, C., & Raatz, U. (1982). Der C-Test: Ein neuer Ansatz zur Messung von allgemeiner Sprachbeherrschung. *AKS Rundbrief 4*, 23-37.

Kos, M., Vosse, T. G., van den Brink, D., & Hagoort, P. (2010). About edible restaurants: Conflicts between syntax and semantics as revealed by ERPs. *Frontiers in Psychology*, *1*, E222.

Kuperberg, G. (2007). Neural mechanisms of language comprehension: challenges to syntax. *Brain Research*, *1146*, 23-49.

Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences*, *4(12)*, 463-470.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: Finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology*, *62*, 14.1-14.22.

Kutas, M., & Hillyard, S. A. (1980a). Reading senseless sentences: brain potentials reflect semantic incongruity. *Science*, *207(4427)*, 203-205.

Kutas, M., & Hillyard, S. A. (1980b). Reading between the lines: event-related brain potentials during natural sentence processing. *Brain and Language*, *11*, 354-373.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature*, *307*, 161-163.

Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (de)constructing the N400. *Nature Reviews Neuroscience*, *9(12)*, 920-933.

Leclercq P., & Edmonds, A. (2014). How to assess L2 proficiency? An overview of proficiency assessment research. In P. Leclerq, A. Edmonds, & H. Hilton (Eds.), *Measuring L2 Proficiency: Perspectives from SLA* (pp. 3-24). Bristol: Multilingual Matters.

Lenneberg, E. (1967). *Biological foundations of language*. New York: John Wiley and Sons.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12(3),* 251-285.

Long, M. H. (2005). Problems with supposed counter-evidence to the critical period hypothesis. *International Review of Applied Linguistics*, *43(3),* 287-317.

Luck, S. J. (2005). An introduction to the event-related potential technique. Cambridge, M.A: MIT Press.

McLaughlin, J., Tanner, D., Pitkanen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning*, *60(2)*, 123-150.

Meisel, J. M. (2011). *First and second language acquisition*. Cambridge: Cambridge UP.

Molfese, D. L., Key, A. P., Maguire, M. J., Dove, G. O., & Molfese, V. J. (2008). Event-related evoked potentials (ERPs) in speech perception. In D. B. Pisoni & R. E. Remez (Eds.), *The handbook of speech perception* (pp. 99-122). Oxford: Blackwell.

Moreno, E. M., Rodriguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, *21(6)*, 477-508.

Moreno, E. M., & Kutas, M. (2005). Processing semantic anomalies in two languages: an electrophysiological exploration in both languages of Spanish-English bilinguals. *Cognitive Brain Research*, *22(2)*, 205-220.

Morgan-Short, K., & Tanner, D. (2014). Event related potentials (ERPs). In J. Jegerski & B. VanPatten (Eds.), *Research methods for second language psycholinguistics* (pp. 127-152). New York: Routledge.

Mueller, J. L. (2005). Mechanisms of auditory sentence comprehension in first and second language: An electrophysiological miniature grammar study. PhD Thesis, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig.

Mueller, J. L. (2009). The influence of lexical familiarity on ERP responses during sentence comprehension in language learners. *Second Language Research*, *25(1)*, 43-76.

Mueller, J. L., Girgsdies, S., & Friederici, A. D. (2008). The impact of a semantic-free second-language training on ERPs responses during case processing. *Neuroscience Letters*, *443(2)*, 77-81.

Mueller, J. L., Hahne, A., Fujii, Y., & Friederici, A. D. (2005). Native and nonnative speakers' processing of a miniature version of Japanese as revealed by ERPs. *Journal of Cognitive Neuroscience*, *17(8)*, 1229-1244.

Mueller, J. L., Hirotani, M., & Friederici, A. D. (2007). ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners. *BMC Neuroscience*, *8*.

Mueller, J. L., Rueschemeyer, S. A., & Friederici, A. D. (2006). Aktivitätsmuster im Gehirn: Unterschiede und Gemeinsamkeiten beim Verstehen von Erst- und Zweitsprache. *Neuroforum*, *2*, 176-184.

Muente, T. F., Heinze, H. J. & Mangun, G. (1993). Dissociation of brain activity related to syntactic and semantic aspects of language. *Journal of Cognitive Neuroscience*, *5(3)*, 335-344.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, *44(1)*, 1-35.

Neville, H.J., Nicol, J., Barss, A., Forster, K., & Garrett, M. (1991). Syntactically based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience*, *3(2),* 155-170.

Newman, A. Tremblay, A., Nichols, E., Neville, H., & Ullman, M. (2012). The influence of language Proficiency on lexical-semantic processing in native and late learners of English: ERP evidence. *Journal of Cognitive Neuroscience*, *24(5)*, 1205-23.

Ojima, S., Nakata, H., & Kagigi, R. (2005). An ERP study of second language learning after childhood: Effects of Proficiency. *Journal of Cognitive Neuroscience*, *17(8)*, 1212-1228.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9(1)*, 97-113.

Osterhout, L., & Holcomb, P. J. (1992). Event-related brain potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31(6)*, 785-806.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Sentences in the brain: Event-related potentials as real-time reflections of sentence comprehension and language learning. In M. Carreiras & C. Clifton, Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP, and beyond* (pp. 271-308). New York: Psychology Press.

Osterhout, L., McLaughlin, J., Pitkanen, I., Frenck-Mestre, C., & Molinaro, N. (2006). Novice learners, longitudinal designs, and event-related potentials: A paradigm for exploring the neurocognition of second-language processing. *Language Learning*, *56(1)*, 199-230.

Osterhout, L., & Mobley, L.A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language*, *34(6)*, 739-773.

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, *14(3)*, 283-317.

Pablos, L., Saddy, J.D., Shirley, E., Williams, N., Erdozia, K.., & Laka, I. (2011). *Licensing of negative polarity items in L1 and L2 speakers of Basque: An ERP study.* Retrieved from: http://www.bangor.ac.uk/bilingualism/seminars/documents/LeticiaPablos3rdOctober2011LicensingofNegativePolarityitemsinL1andL2speakersofBasqueanERPst.pdf.

Pakulak, E., & Neville, H. J. (2010). Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials. *Journal of Cognitive Neuroscience*, *22*, 2728-2744.

Papadopoulou, D. & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition*, *25*, 501-528.

Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms.* Princeton: Princeton University Press.

van Petten, C., & Luka, B. J. (2006). Neural localization of semantic context effects in electromagnetic and hemodynamic studies. *Brain and Language*, *97(3)*, 279-293.

R Development Core Team (2014). R: a language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, *114(3)*, 510-532.

Resmeth (2012, November 28). My variables have an unhealthy relationship [Web log comment], retrieved from http://blogs.lse.ac.uk/methodology/2012/11/28/unhealthy/.

Rispens, J., & Krikhaar, E. (2010). Using event-related potentials in language acquisition research. In E. Blom & S. Unsworth (Eds.). *Experimental methods in language acquisition research* (pp. 95-114). Amsterdam: John Benjamins.

Saddy, D. J., Drenhaus H., & Frisch, S. (2004). Processing polarity items: contrastive licensing costs. *Brain and Language*, *90(1)*, 495-502.

Schuette, J. (2006). On the Processing of Negative Polarity Constructions: When the Licensor Makes the Difference. *Proceedings of the First Central European Students Conference in Linguistics at Budapest.* May 24-26. http://www.nytud.hu/cescl/proceedings.html

Sharbrough, F., Chartrian, G. E., Lesser, R. P., Lüders, H., Nuwer, M., & Picton, T. W. (1995). American electroencephalographic society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, *8*, 200-202.

Singleton, D. (2005). The critical period hypothesis: A coat of many colours. *International Review of Applied Linguistics*, *43(4)*, 269-285.

Steinhauer, K. (2014). Event-related potentials (ERPs) in second language research: A brief introduction to the technique, a selected review, and an invitation to reconsider critical periods in L2. *Applied Linguistics*, *35(4)*, 393-417.

Steinhauer, K., & Connolly, J. F. (2008). Event-related potentials in the study of language. In B. Stemmer & H. Whitaker (Eds.), *Handbook of the cognitive neuroscience of language* (pp. 91-104). New York: Elsevier.

Steinhauer, K., Drury, J. E., Portner, P., Walenski, M., & Ullman, M. T. (2010). Syntax, concepts, and logic in the temporal dynamics of language comprehension: Evidence from event-related potentials. *Neuropsychologia*, *48(6)*, 1525-1542.

Steinhauer, K., White, E., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related brain potentials. *Second Language Research*, *25(1)*, 13-41.

Tanner, D., Inoue, K., & Osterhout, L. (2014). Brain-based individual differences in on-line L2 sentence comprehension. *Bilingualism: Language and Cognition*, *17*, 277-293.

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition: ERP evidence. *Bilingualism: Language and Cognition*, *16(2)*, 367-382.

Thomas, M. (1994). Assessment of L2 Proficiency in second language acquisition research. *Language Learning, 44(2)*, 307-336.

Tremblay, A., Dalhousie University, Ransijn, J. and University of Copenhagen (2013). LMERConvenienceFunctions: A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions. R package version 2.5. URL:http://CRAN.R-project.org/package=LMERConvenienceFunc-tions.

Vanhove, J. (2013). The critical period hypothesis in second language acquisition: A statistical critique and a reanalysis. *PLOS ONE*, *8*. e69172.

Weber-Fox, C. M., & Neville,H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8(3)*, 231-256.

van der Wouden, T. (1997). Negative contexts. Collocation, polarity and multiple negation. New York: Routledge.

Wurm, L. H., & Fisicaro, S. A. (2014). What residualizing predictors in regression analyses does (and what it does not do). *Journal of Memory and Language*, 72, 37-48.

Ye, Z., Zhan, W., & Zhou, X. (2007). The semantic processing of syntactic structure in sentence comprehension: An ERP study. *Brain Research*, *1142*, 135-145.

Yurchenko, A., den Ouden, D. B., Hoeksema, J., Dragov, O., Hoeks, J. C. J., & Stowe, L. A. (2013). Processing polarity: ERP evidence for differences between positive and negative polarity. *Neuropsychologia*, *51(1)*, 132-141.

Zschocke, S., & Hansen, H.-C. (Eds) (2011). *Klinische Enzephalographie* (2.Aufl.). Berlin: Springer.

List of Abbreviations

ACC – Accusative

AoA – Age of Acquisition

CEFR – Common European Framework of Reference for Language

Chap. – Chapter

df – Degree(s) of Freedom

Diff – Difference

EAHP – Early AoA and High Proficiency

EEG – Electroencephalogram

EOG – Electroocolugram

ERP – Event-Related Potential

Exp. – Example

Fig. – Figure

HEOG – Horizontal Electroocolugram

L1 – Native Language

L2 – Second Language

LAHP – L2 Learners with Late AoA and High Proficiency

LALP – L2 Learners with Late AoA and Low Proficiency

min – Minute

ms – Millisecond

μV – Micro Volt(s)

NOM – Nominative

NP – Nominal Phrase

NPI – Negative Polarity Item

PPI – Positive Polarity Item

ROI – Region of Interest

RT – Reaction Time

V - Verb

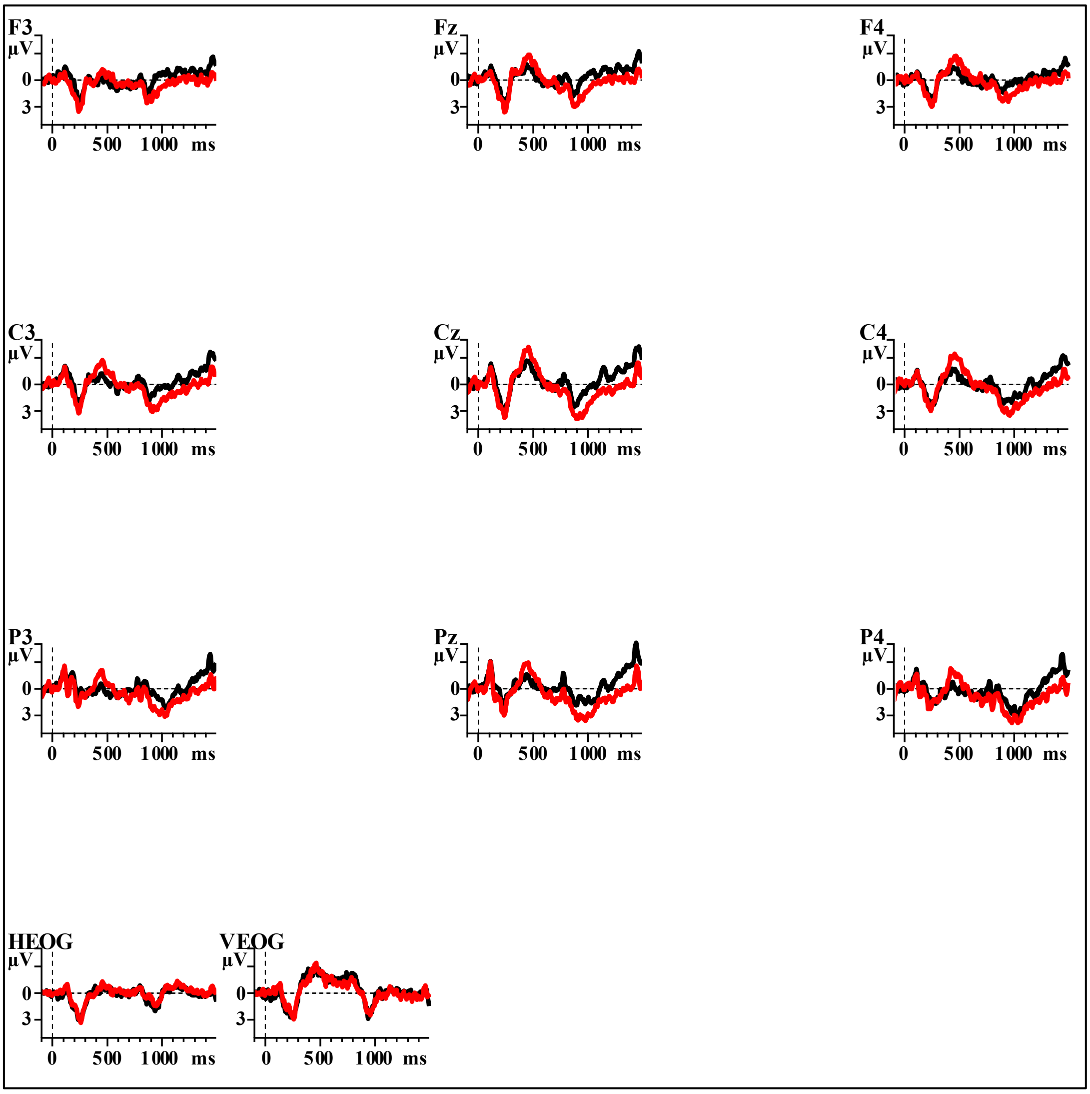
VEOG – Vertical Electroocolugram

List of Tables

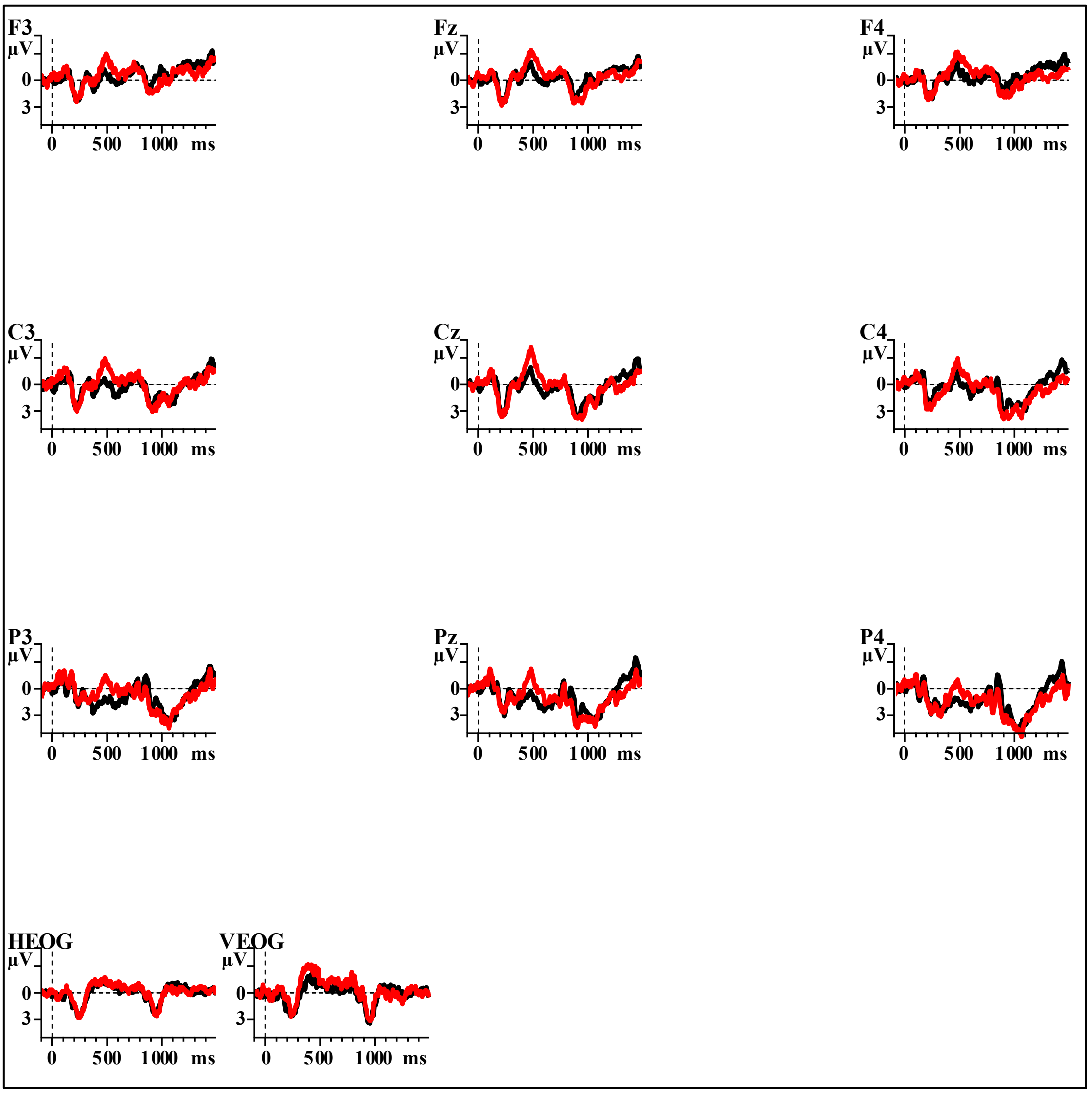
List of Figures

1 Appendix 1: ERP Figures of Single Groups

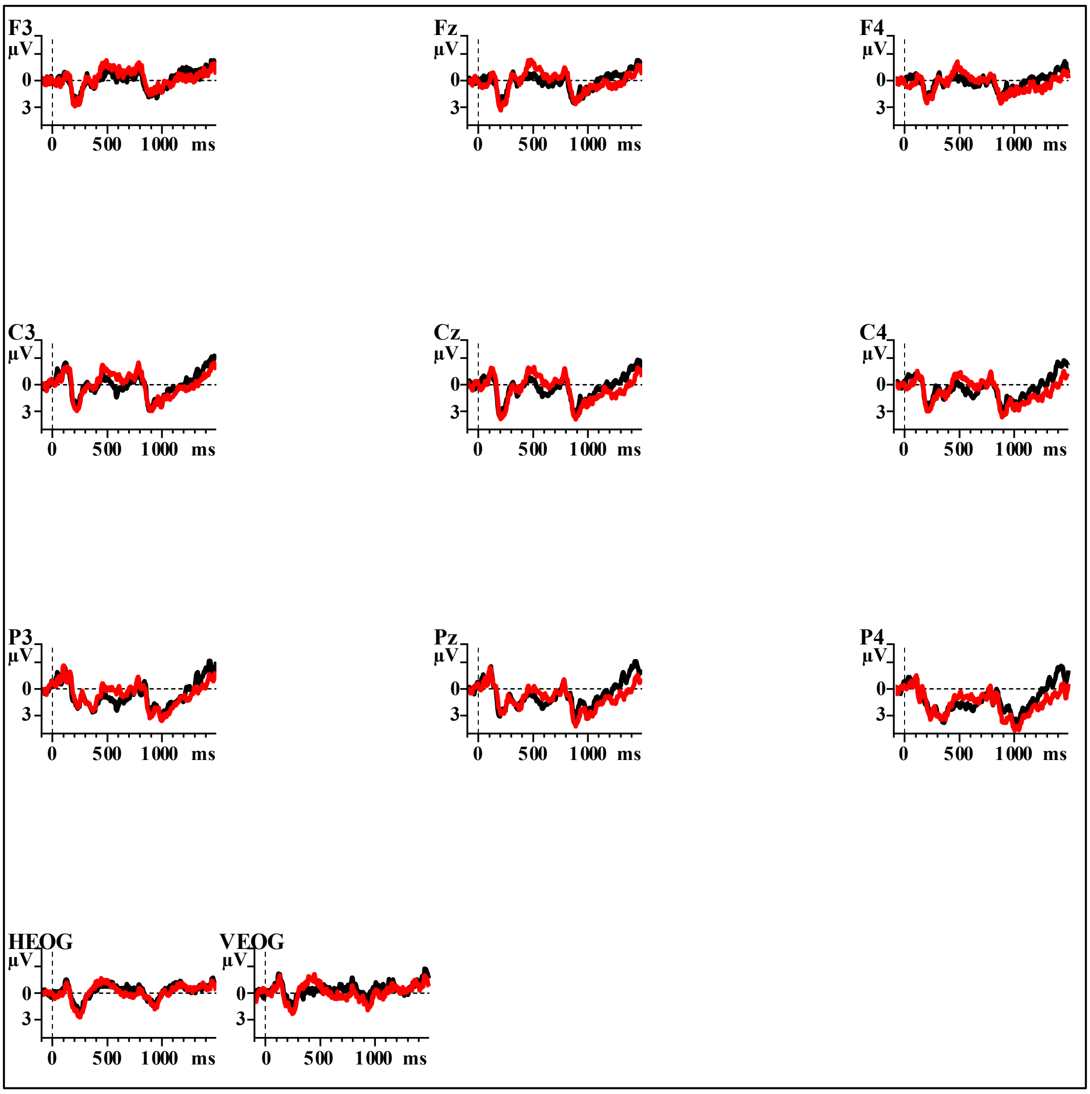
1.1 Semantic Incongruity



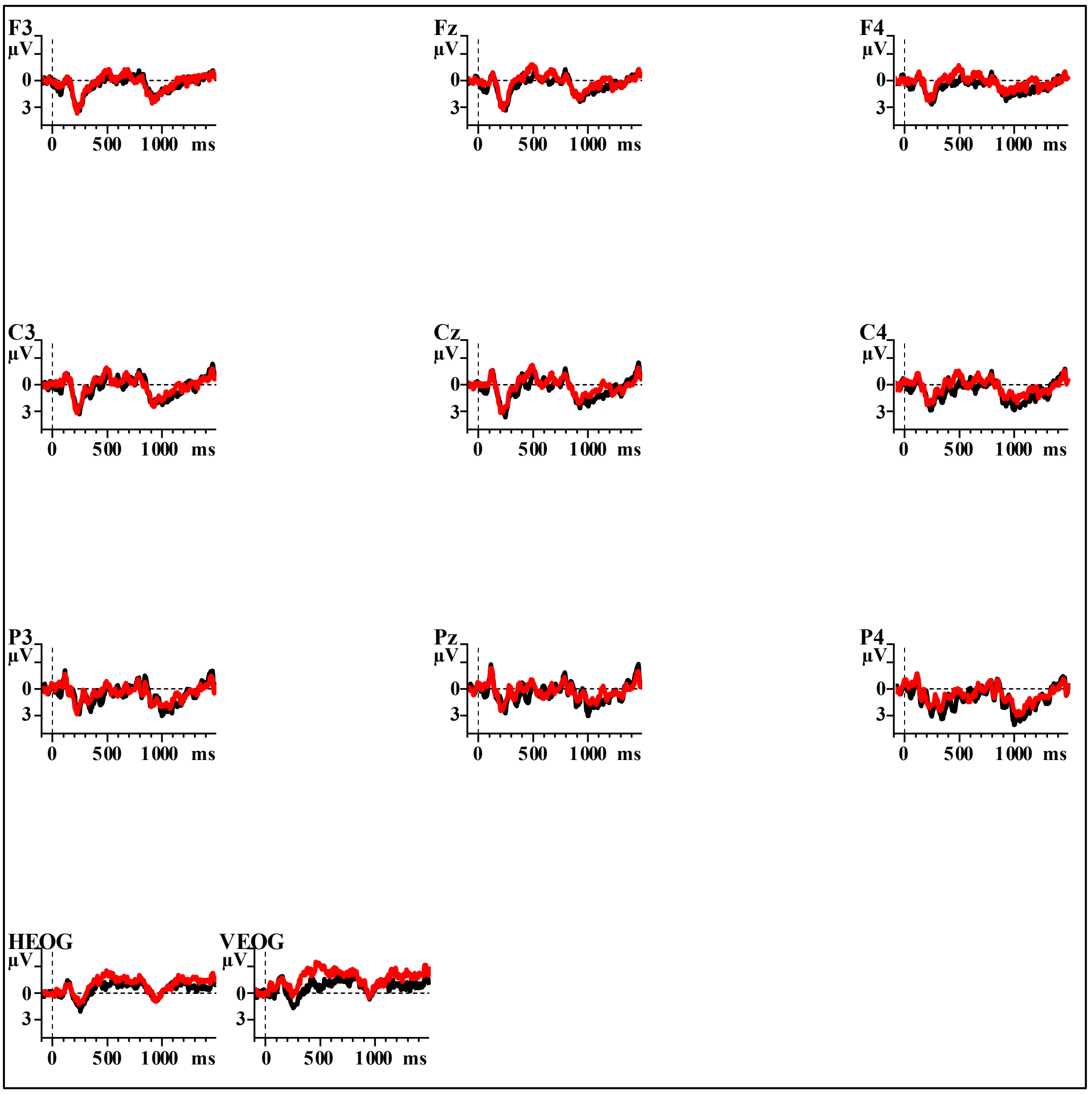
Appendix Figure 1.1: Grand average ERPs according to the Congruity conditions (black = congruent; red = incongruent) elicited by native speakers. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.



Appendix Figure 1.2: Grand average ERPs according to the Congruity conditions (black = congruent; red = incongruent) elicited by EAHP. Voltages are plotted on y-axis ranging from -4μV to 4μV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

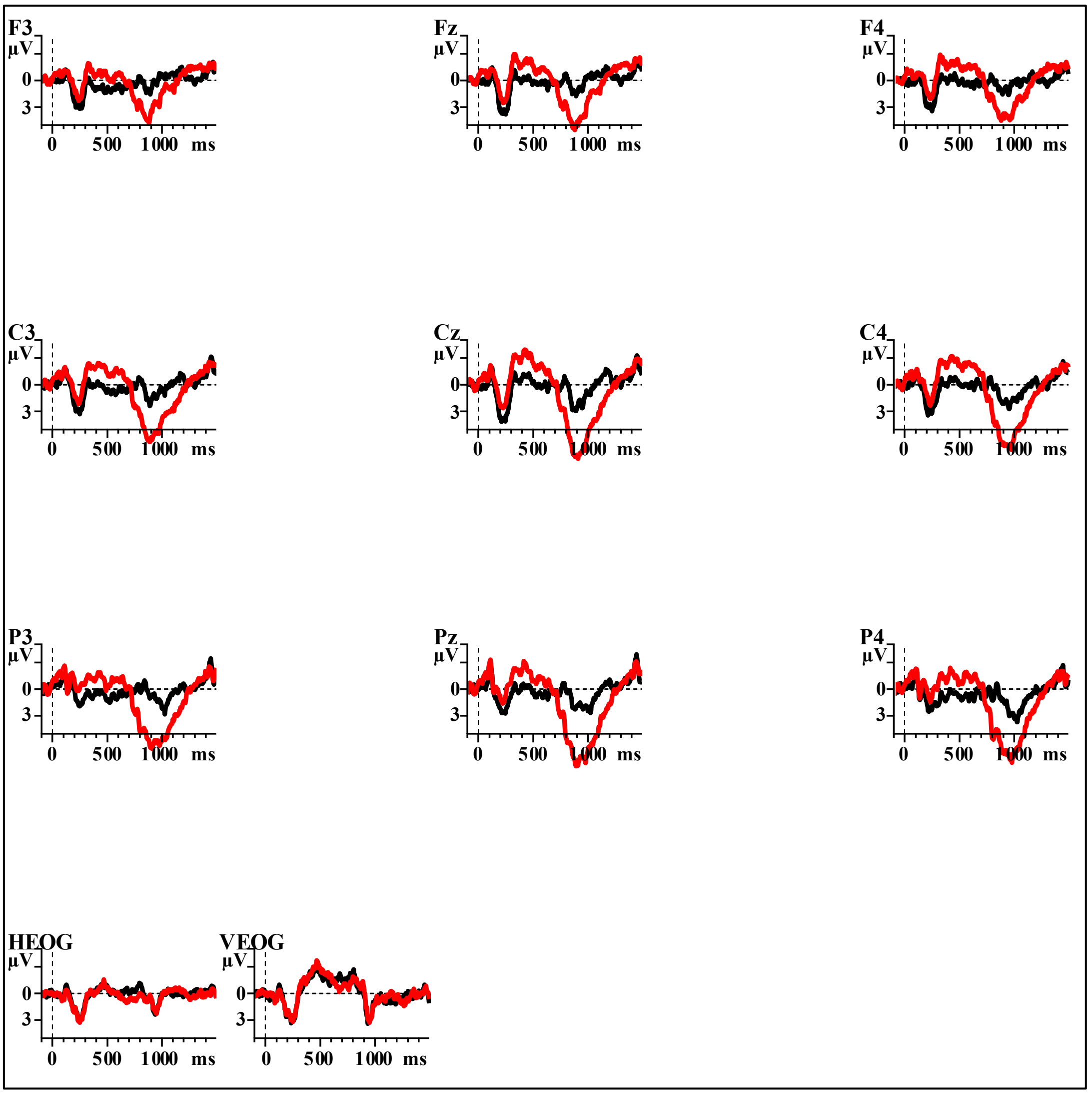


Appendix Figure 1.3: Grand average ERPs according to the Congruity conditions (black = congruent; red = incongruent) elicited by LAHP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

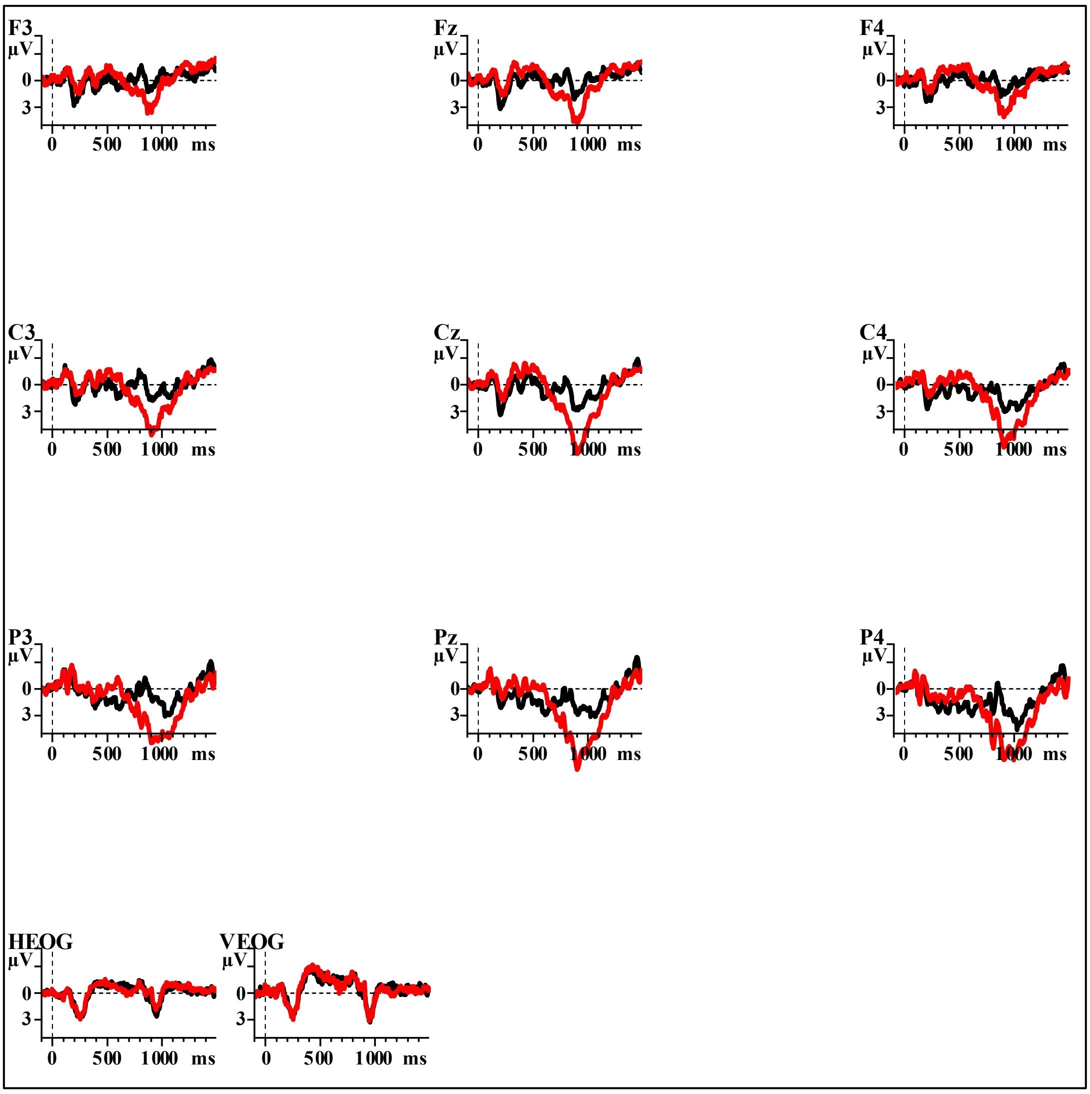


Appendix Figure 1.4: Grand average ERPs according to the Congruity conditions (black = congruent; red = incongruent) elicited by LALP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

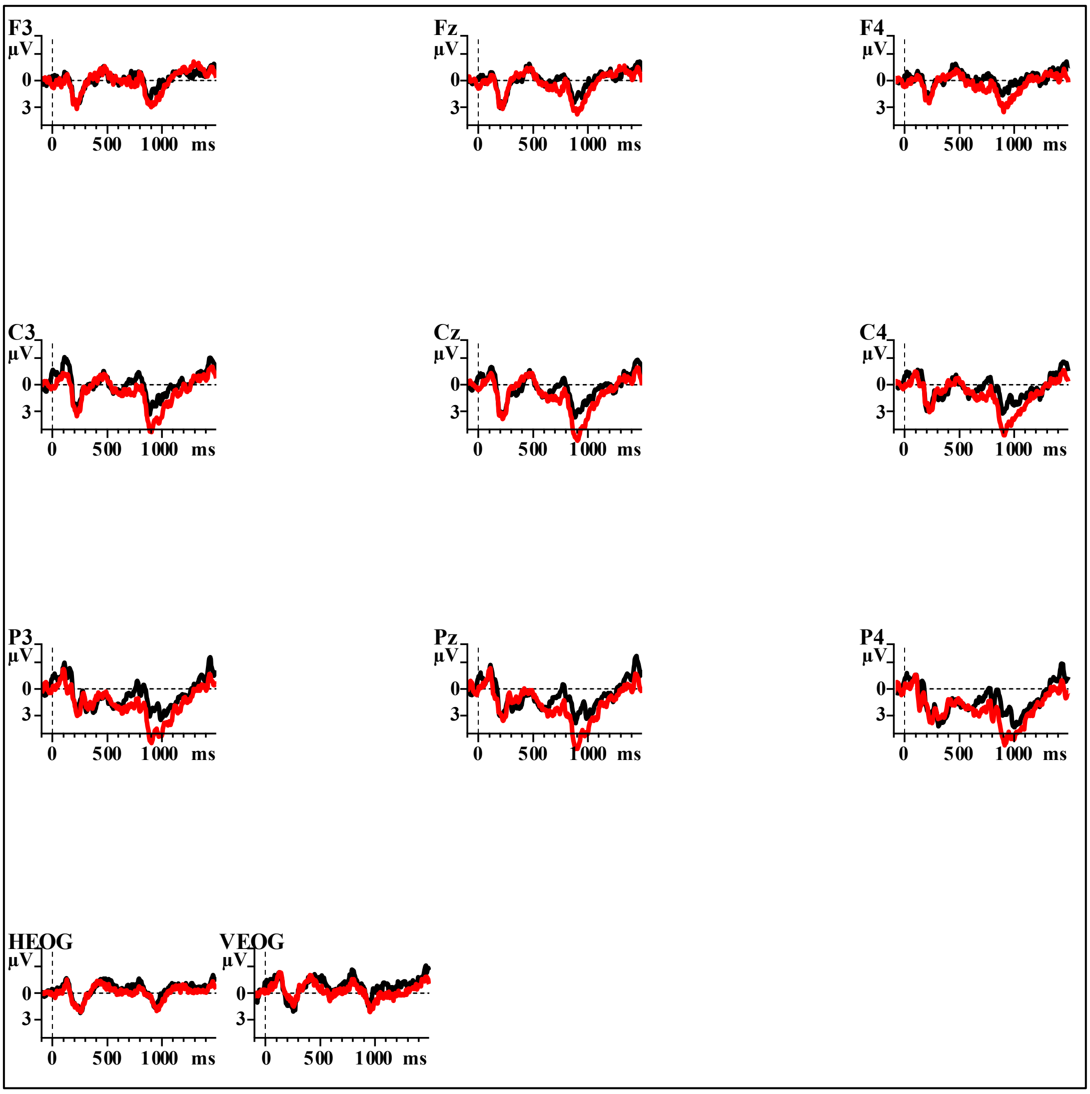
1.2 Double Nominative Violation



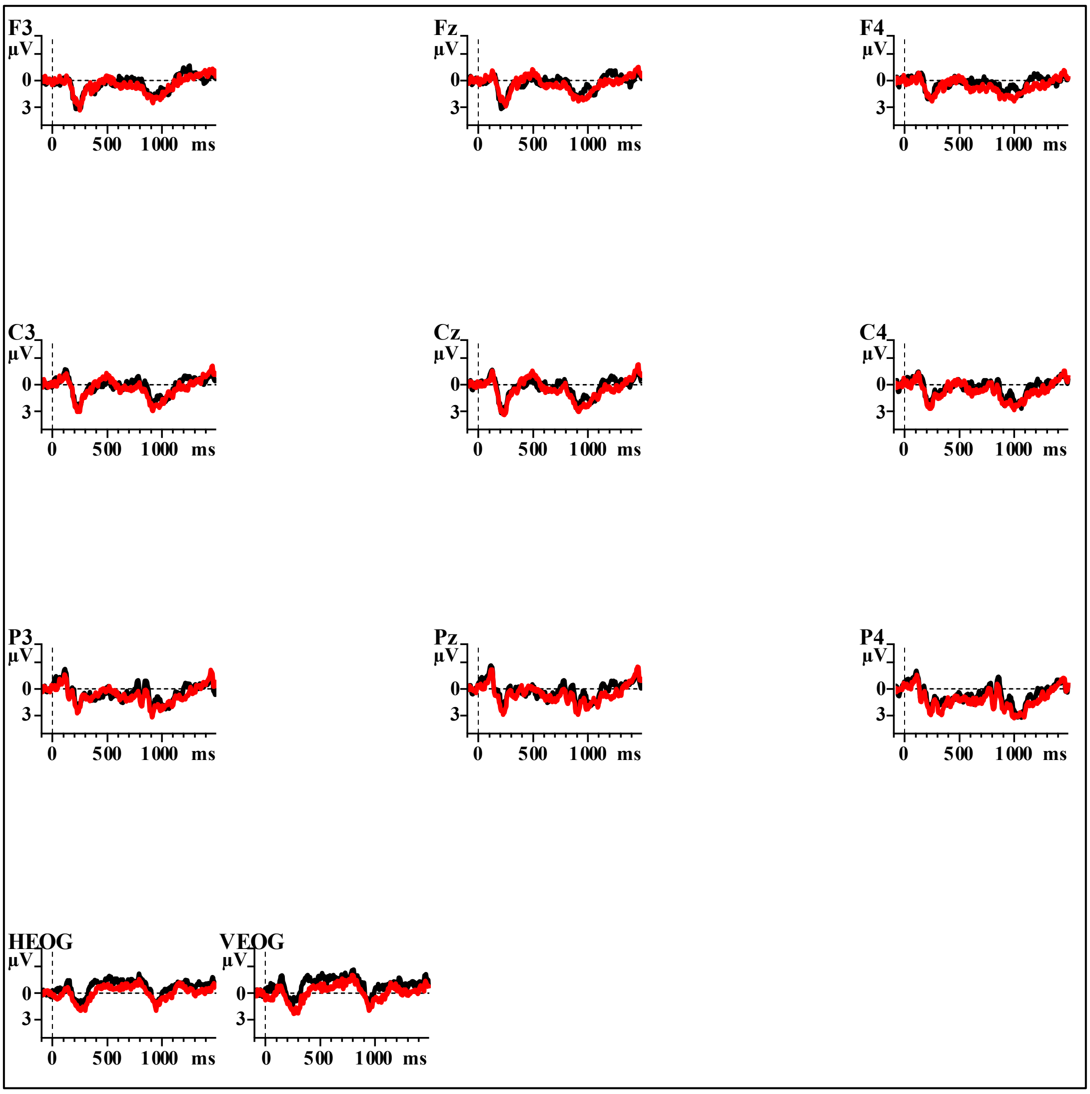
Appendix Figure 1.5: Grand average ERPs according to the Case conditions (black = case-congruent; red = case-violation) elicited by native speakers. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.



Appendix Figure 1.6: Grand average ERPs according to the Case conditions (black = case-congruent; red = case-violation) elicited by EAHP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

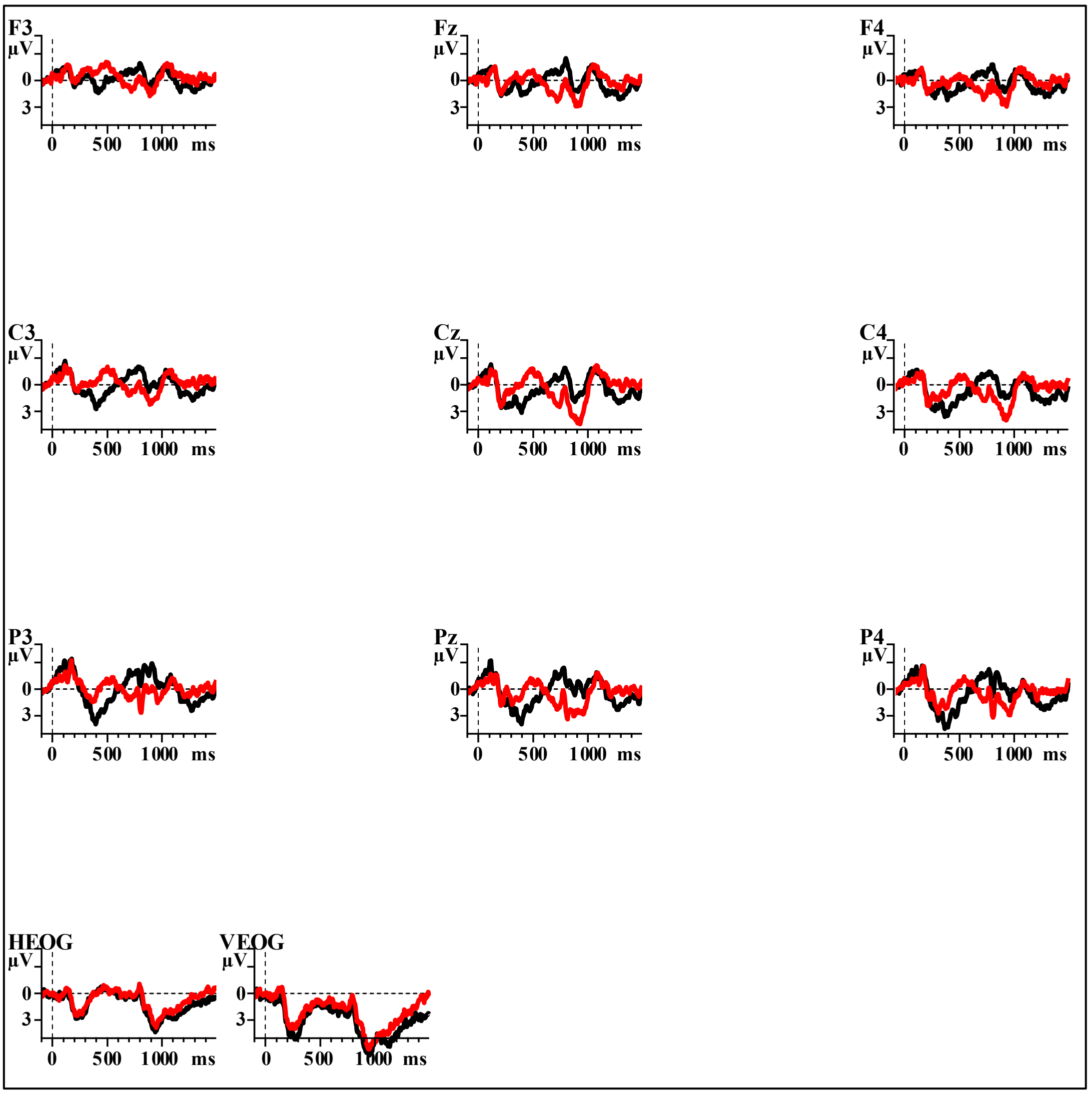


Appendix Figure 1.7: Grand average ERPs according to the Case conditions (black = case-congruent; red = case-violation) elicited by LAHP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

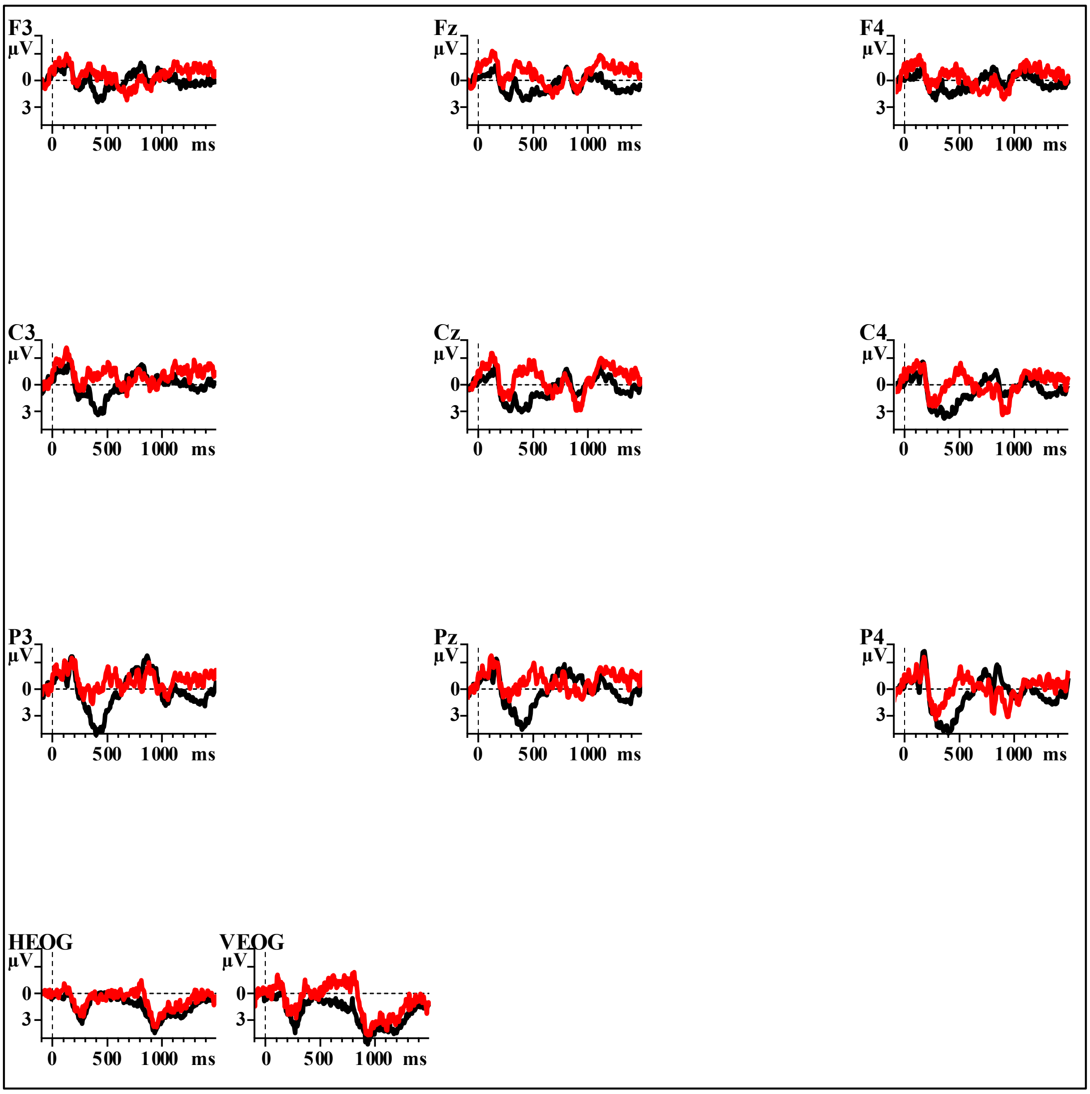


Appendix Figure 1.8: Grand average ERPs according to the Case conditions (black = case-congruent; red = case-violation) elicited by LALP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

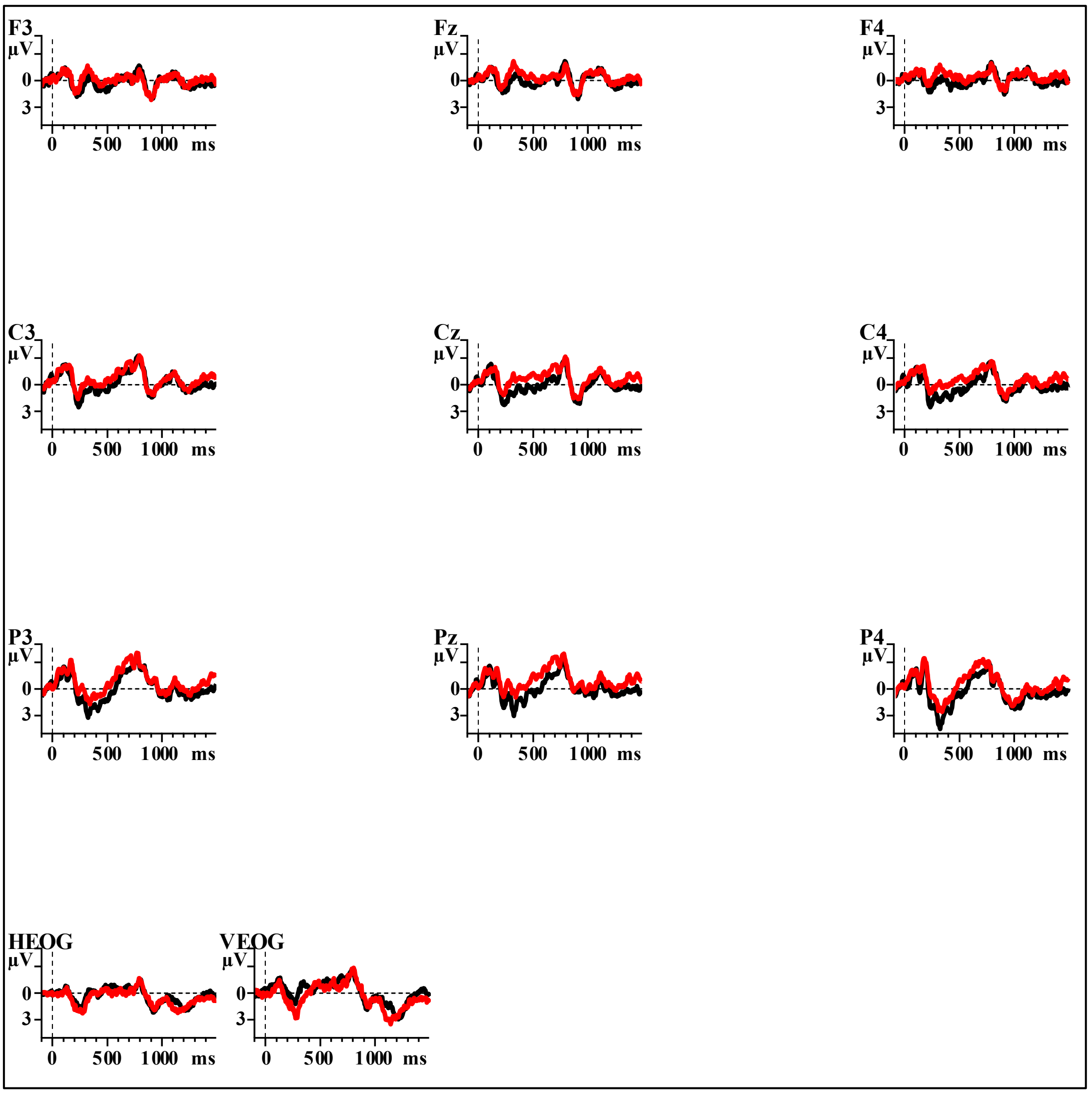
1.3 NPI Licensing



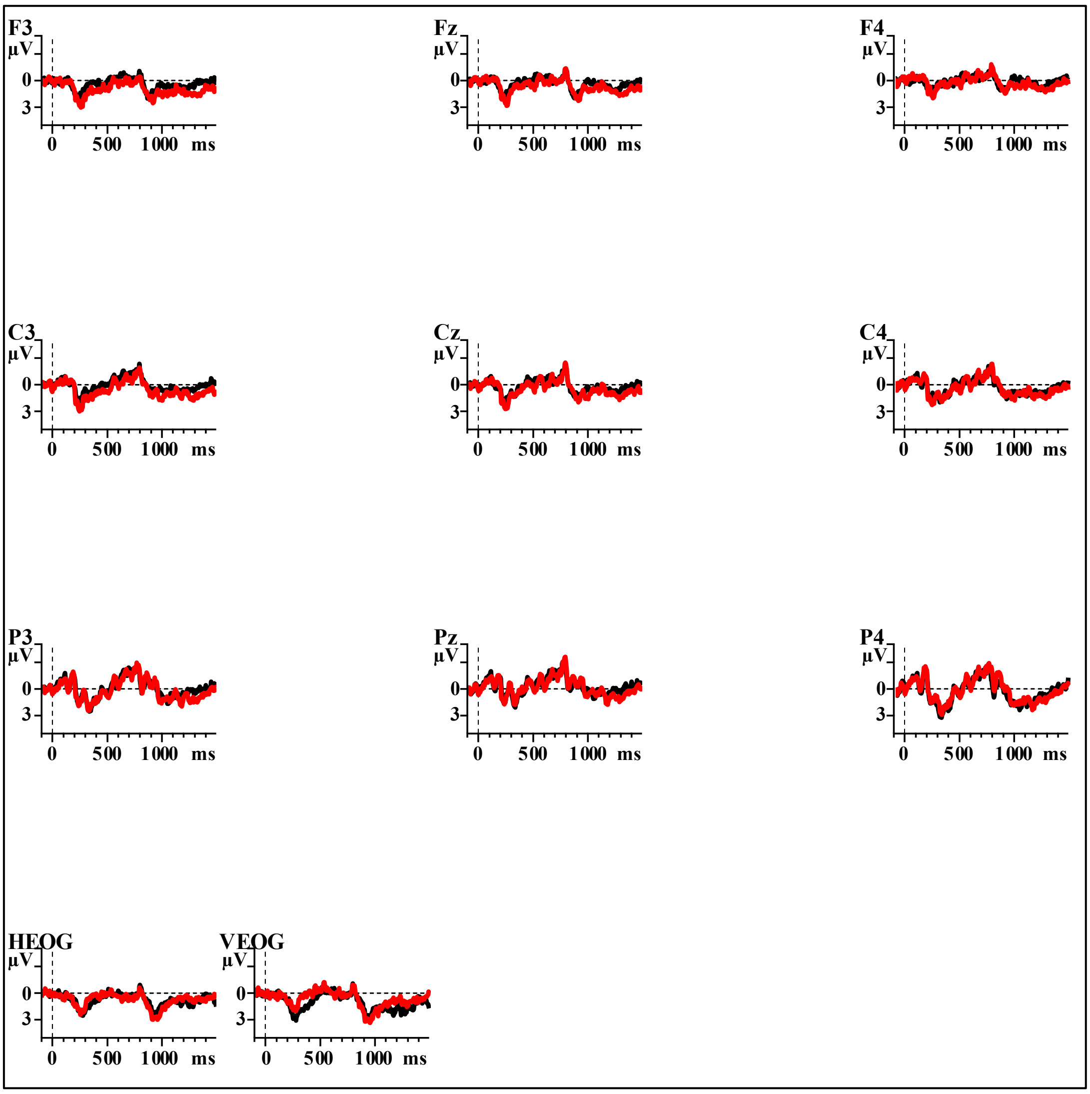
Appendix Figure 1.9: Grand average ERPs according to the Licensor conditions (black = licensed; red = non-licensed) elicited by native speakers. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.



Appendix Figure 1.10: Grand average ERPs according to the Licensor conditions (black = licensed; red = non-licensed) elicited by EAHP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.



Appendix Figure 1.11: Grand average ERPs according to the Licensor conditions (black = licensed; red = non-licensed) elicited by LAHP. Voltages are plotted on y-axis ranging from -4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.



Appendix Figure 1.12: Grand average ERPs according to the Licensor conditions (black = licensed; red = non-licensed) elicited by LALP. Voltages are plotted on y-axis ranging from - 4µV to 4µV. Time array is plotted on x-axis and ranges from -100 to 1500 ms. Stimulus onset occurred at 0 ms. Selected electrodes are labelled accordingly. HEOG = Horizontal Electroocolugram. VEOG = Vertical Electroocolugram. Negative voltages are plotted up.

2 Appendix: Statistical Tables Experiment 1

Time-window 400–500 ms

Native Speakers: Statistical results of single ROI analysis (due to planned comparisons) for the Congruity effect in each ROI are listed in Appendix Table 2-1.

Appendix Table 2-1: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for native speakers. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	142/106	-0.67	1	5.21	5.21	1.44	.23	.23
MIDLINE-ANTERIOR*	106/70	-1.15	1	2.14	2.14	2.13	.14	.14
RIGHT-ANTERIOR*	142/106	-1.17	1	6.09	6.09	3.87	.05	.05
LEFT-POSTERIOR*	142/106	-1.13	1	5.28	5.28	3.11	.08	.08
MIDLINE-POSTERIOR*	106/70	-1.33	1	5.89	5.89	2.75	.09	.09
RIGHT-POSTERIOR	137/101	-1.42	1	4.51	4.51	5.05	.03	.03

Time-window 400–425 ms

Native Speakers: Statistical results for native speakers' ERPs in the 400–425 ms time-window only reveals a main effect for ROI ($F(5,761) = 14.53$, $p < .001$)[121]. Results of single ROI analysis (due to planned comparisons) for the Congruity effect in each ROI are summarised in Appendix Table 2-2.

[121] *Denominator lower bound* = 599, $p < .001$.

Appendix Table 2-2: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 400–425 ms conducted for native speakers. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in µV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	142/106	-0.72	1	8.39	8.39	2.13	.14	.14
MIDLINE-ANTERIOR*	106/70	-1.15	1	3.46	3.46	2.54	.11	.11
RIGHT-ANTERIOR	140/104	-1.03	1	7.02	7.02	4.39	.04	.04
LEFT-POSTERIOR*	138/106	-1.24	1	5.27	5.27	4	.05	.05
MIDLINE-POSTERIOR*	106/70	-1.21	1	6.33	6.33	2.32	.13	.13
RIGHT-POSTERIOR*	142/106	-1.33	1	8.91	8.91	3.84	.05	.05

L2 learners: Statistical results for L2 learners' ERPs in the 400–425 ms time-window only reveals a main effect for ROI ($F(5,2575) = 18.74$, $p < .001$) [122]. Results of single ROI analysis (due to planned comparisons) for the Congruity effect in each ROI are summarised in Appendix Table 2-3.

Appendix Table 2-3: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 400–425 ms conducted for L2 learners. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in µV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	478/358	-0.15	1	0.46	0.46	< 1		
MIDLINE-ANTERIOR*	358/238	-0.36	1	0.52	0.52	< 1		
RIGHT-ANTERIOR*	478/358	-0.17	1	0.61	0.61	< 1		
LEFT-POSTERIOR*	478/358	-0.49	1	2.07	2.07	1.47	.22	.22
MIDLINE-POSTERIOR*	358/238	-0.27	1	0.36	0.36	< 1		
RIGHT-POSTERIOR*	478/358	-0.2	1	0.58	0.58	< 1		

[122] *Denominator lower bound* = 2035, $p < .001$.

Time-window 425–450 ms

Native Speakers: Statistical results for native speakers' ERPs in the 425–450 ms time-window only reveals a main effect for ROI ($F(5{,}764) = 24.66$, $p < .001$) [123]. Results of single ROI analysis due to planned comparisons for the Congruity effect in each ROI are summarised in Appendix Table 2-4.

Appendix Table 2-4: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 425–450 ms conducted for native speakers. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	142/106	-0.51	1	2.58	2.58	< 1		
MIDLINE-ANTERIOR*	106/70	-1.07	1	1.46	1.46	1.1	.29	.29
RIGHT-ANTERIOR*	142/106	-1.06	1	4.04	4.04	2.28	.13	.13
LEFT-POSTERIOR*	142/106	-1.1	1	3.79	3.79	1.94	.16	.16
MIDLINE-POSTERIOR*	106/70	-1.31	1	3.99	3.99	1.68	.19	.19
RIGHT-POSTERIOR*	142/106	-1.34	1	4.3	4.3	2.28	.13	.13

L2 learners: Statistical results for L2 learners' ERPs in the 425–450 ms time-window as shown in Appendix Table 2-5 do not reveal any significant main effects for AoA, proficiency nor any reliable term of higher order involving both factors. Statistical results according to the significant differences of the Congruity effect with respect to its distribution is given in Appendix Table 2-6.

[123] *Denominator lower bound* = 602, $p < .001$.

Appendix Table 2-5: ANOVA table of the average ERP amplitudes in the time-window 425–450 ms conducted for L2 learners according to the Congruity conditions (*denominator upper bound df* = 2565, *denominator lower bound df* = 2085).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CONGRUITY	1	9.58	9.58	7.56	.01	.01
ROI	5	143.47	28.69	22.64	< .001	< .001
CONGRUITY×ROI	5	16.49	3.3	2.6	.02	.02

Appendix Table 2-6: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 425–450 ms conducted for L2 learners. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	478/358	-0.43	1	5.04	5.04	2.23	.13	.13
MIDLINE-ANTERIOR	349/229	-0.92	1	3.3	3.3	628	.01	.01
RIGHT-ANTERIOR	467/347	-0.66	1	4.75	4.75	4.91	.03	.03
LEFT-POSTERIOR	469/349	-0.91	1	7.62	7.62	7.02	.01	.01
MIDLINE-POSTERIOR	350/230	-1	1	4.89	4.89	5.7	.02	.02
RIGHT-POSTERIOR	466/346	-1	1	8.36	8.36	7.53	.01	.01

Time-window 450–475 ms

Native Speakers: Statistical results for native speakers' ERPs in the 450–475 ms time-window only reveals a main effect for ROI ($F(5,763) = 39.6$, $p < .001$) [124]. Results of single ROI analysis (due to planned comparisons) for the Congruity effect in each ROI are summarised in Appendix Table 2-7.

[124] *Denominator lower bound* = 601, $p < .001$.

Appendix Table 2-7: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 450–475 ms conducted for native speakers. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	142/106	-0.98	1	10.28	10.28	2.37	.12	.12
MIDLINE-ANTERIOR*	106/70	-1.44	1	3.34	3.34	2.77	.09	.09
RIGHT-ANTERIOR	137/101	-1.33	1	3.45	3.45	4	.05	.05
LEFT-POSTERIOR*	142/106	-1.15	1	5.32	5.32	2.85	.09	.09
MIDLINE-POSTERIOR*	106/70	-1.4	1	6.21	6.21	2.64	.1	.1
RIGHT-POSTERIOR*	142/106	-1.33	1	6.86	6.86	3.5	.06	.06

L2 learners: Statistical results for L2 learners' ERPs in the 450–475 ms time-window as shown in Appendix Table 2-8 do not reveal any significant main effects for AoA, proficiency nor any reliable influence on the strength of the Congruity effect. Yet with increasing AoA mean potentials become more negative on left-posterior ROIs. Statistical results according to the significant differences of the Congruity effect with respect to its distribution is given in Appendix Table 2-9. It appears numerically largest on midline ROIs.

Appendix Table 2-8: ANOVA table of the average ERP amplitudes in the time-window 450–475 ms conducted for L2 learners according to the Congruity conditions (*denominator upper bound df* = 2546, *denominator lower bound df* = 2066).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CONGRUITY	1	15.87	15.87	12.93	<.001	<.001
ROI	5	159.13	31.83	25.93	< .001	< .001
AOA	1	6.17	6.17	5.02	.03	.03
CONGRUITY×ROI	5	47.38	9.48	7.72	<.001	<.001
ROI×AOA	5	14.62	2.92	2.38	.04	.04

Appendix Table 2-9: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 450–475 ms conducted for L2 learners. Diff = Difference.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	460/340	-0.7	1	8.93	8.93	6.05	.01	.01
MIDLINE-ANTERIOR	349/229	-1.57	1	7	7	15.06	<.001	<.001
RIGHT-ANTERIOR	463/343	-0.9	1	7.36	7.36	8.23	<.001	<.001
LEFT-POSTERIOR	472/352	-1.33	1	12.6	12.6	12.36	<.001	<.001
MIDLINE-POSTERIOR	350/230	-1.51	1	10.95	10.95	12.25	<.001	<.001
RIGHT-POSTERIOR	468/348	-1.17	1	15.03	15.03	11.81	<.001	<.001

Time-window 475–500 ms

Native Speakers: Statistical results for native speakers' ERPs in the 475–500 ms are listed in Appendix Table 2-10. There is a main effect for Congruity indicating a significant negativity effect which further is differently distributed. Results of single ROI analysis for the Congruity effect in each ROI are summarised in Appendix Table 2-11. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

Appendix Table 2-10: ANOVA table of the average ERP amplitudes in the time-window 475–500 ms conducted for native speakers according to the Congruity conditions (*denominator upper bound df* = 756, *denominator lower bound df* = 612).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CONGRUITY	1	4.93	4.93	3.87	.05	.05
ROI	5	178.1	35.6	27.95	< .001	< .001
CONGRUITY×ROI	5	26.11	5.22	4.1	<.001	<.001

Appendix Table 2-11: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 475–500 ms conducted for native speakers. Diff = Difference. *-indexed ROIs list statistical values taken from the original most complex model. The refitted model only had an intercept.

ROI	*denominators upper-/lower-bound*	*mean Diff in µV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR*	142/106	-0.48	1	1.92	1.92	< 1		
MIDLINE-ANTERIOR*	106/70	-0.96	1	1.16	1.16	1.57	.21	.21
RIGHT-ANTERIOR	137/101	-1.26	1	5.13	5.13	4.67	.03	.03
LEFT-POSTERIOR*	142/106	-1.03	1	4.49	4.49	2.67	.1	.1
MIDLINE-POSTERIOR*	106/70	-1.39	1	7.37	7.37	4.09	.05	.05
RIGHT-POSTERIOR	137/101	-1.69	1	7.74	7.74	9.31	<.001	<.001

L2 learners: Statistical results for L2 learners' ERPs in the 475–500 ms time-window are shown in Appendix Table 2-12. Results do not reveal any significant main effects for AoA, proficiency nor any reliable influence on the strength of the Congruity effect. Yet, again, with increasing AoA mean potentials become more positive on midline- and right-anterior ROIs. Statistical results according to the significant differences of the Congruity effect with respect to its distribution is given in Appendix Table 2-13. It appears numerically largest on midline-anterior electrodes.

Appendix Table 2-12: ANOVA table of the average ERP amplitudes in the time-window 475–500 ms conducted for L2 learners according to the Congruity conditions (*denominator upper bound df* = 2549, *denominator lower bound df* = 2069).

COEFFICIENTS	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p*	*lower p*
CONGRUITY	1	14.56	14.56	12.66	<.001	<.001
ROI	5	194.92	38.98	33.91	< .001	< .001
AOA	1	3.82	3.82	3.32	.07	.07
CONGRUITY×ROI	5	34.37	6.87	5.98	<.001	<.001
ROI×AOA	5	21.65	4.33	3.77	<.001	<.001

Appendix Table 2-13: ANOVA table of the Congruity effect in each ROI according to the average ERP amplitudes in the time-window 475–500 ms conducted for L2 learners. Diff = Difference.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	466/346	-1	1	10.89	10.89	7.94	.01	.01
MIDLINE-ANTERIOR	347/227	-1.66	1	6.62	6.62	15.21	<.001	<.001
RIGHT-ANTERIOR	463/343	-1.04	1	13.18	13.18	16.24	<.001	<.001
LEFT-POSTERIOR	466/346	-1.14	1	7.28	7.28	8.4	<.001	<.001
MIDLINE-POSTERIOR	347/227	-1.25	1	8.37	8.37	9.53	<.001	<.001
RIGHT-POSTERIOR	466/346	-0.96	1	8.3	8.3	7.73	.01	.01

3 Appendix: Statistical Tables Experiment 2

Time-window 400–500 ms

Native Speakers: Statistical results of single ROI analysis (due to planned comparisons) for the Case effect in each ROI are listed in Appendix Table 3-1. The negativity effect is numerically larger on midline than lateral ROIs.

Appendix Table 3-1: ANOVA table of the Case effect in each ROI according to the average ERP amplitudes in the time-window 400–500 ms conducted for native speakers. Diff = difference.

ROI	*denominators upper-/lower-bound*	*mean Diff in μV*	*df*	*sum Sq*	*mean Sq*	*F-value*	*upper p-value*	*lower p-value*
LEFT-ANTERIOR	136/100	-1.56	1	25.8	25.8	13.72	<.001	<.001
MIDLINE-ANTERIOR	103/67	-2.29	1	10.35	10.35	13.11	<.001	<.001
RIGHT-ANTERIOR	137/101	-1.58	1	9.29	9.29	9.16	<.001	<.001
LEFT-POSTERIOR	136/100	-1.58	1	10.75	10.75	11.36	<.001	<.001
MIDLINE-POSTERIOR	102/66	-1.92	1	12.28	12.28	11.27	<.001	<.001
RIGHT-POSTERIOR	138/102	-1.7	1	12.63	12.63	11.79	<.001	<.001

4 Appendix: Structures of Most Complex Statistical Models

This appendix lists the random- and mixed-effects structures of the most complex statistical models performed on the behavioural and ERP data-sets as they were implemented in R.

4.1 Accuracy: Generalised Mixed-Effects Model

Native speakers: model <- glmer(Accuracy ~ Condition + (1+Condition|Subject) + (1|Item),

data,family="binomial")

L2 learners: model <- glmer(Accuracy ~ (Condition*AoA*Proficiency) +

(1+Condition|Subject) + (1|Item), data, family="binomial")

4.2 Reaction Times: Linear Mixed-Effects Model

Native speakers: model <- lmer(RT.log ~ Condition + (1+Condition|Subject) + (1|Item),

data)

L2 learners: model <- lmer(RT.log ~ (Condition*AoA*Proficiency) +

(1+Condition|Subject) + (1|Item), data)

4.3 ERPs: Linear Mixed-Effects Model

Native Speakers: model <- lmer(Potential ~ (Condition*ROI) + (1+Condition|Subject) +

(0+ROI|Subject), data)

L2 learners: model <- lmer(Potential ~ (Condition+ROI+AoA+Proficiency)^3 +

(1+Condition|Subject) + (0+ROI|Subject), data)

5 Appendix: Stimulus Sentences

5.1 Experiment 1: Semantic Incongruity

Der / Kein Tiger frisst den Knochen / *Schatten und bekommt einen Krampf.

Der / Kein Junge klaut den Fernseher / *Ausgang und bekommt einen Klaps.

Der / Kein Chef beurlaubt den Vorstand / *Tisch und entlässt einen Hausmeister.

Der / Kein Onkel verschenkt den Lolli / *Unfall und umarmt einen Freund.

Der / Kein Ritter verteidigt den Besitz / *Bildschirm und bekommt einen Kuss.

Der / Kein Pförtner zerreißt den Vertrag / *Husten und öffnet einen Eingang.

Der / Kein Lehrer fährt den Wagen / *Park und hat einen Unfall.

Der / Kein Maurer singt den Song / *Computer und bekommt einen Vertrag.

Der / Kein Opa zerschlägt den Krug / *Saft und kauft einen Becher.

Der / Kein Hund beißt den Fuß / *Traum und bekommt einen Klaps.

Der / Kein Schmied biegt den Ring / *Käse und verkauft einen Säbel.

Der / Kein Förster fällt den Baum / *Roller und rettet einen Hirsch.

Der / Kein Passant löscht den Brand / *Tag und bekommt einen Preis.

Der / Kein Schüler erreicht den Abschluss / *Stiefel und bekommt einen Preis.

Der / Kein Stürmer schießt den Freistoß / *Zettel und erzielt einen Treffer.

Kein Student liest den Roman / *Mülleimer und schreibt einen Aufsatz.

Der / Kein Professor überlegt den Satz / *Schrank und löst ein Problem.

Der / Kein Hund findet den Unterschlupf / *Druck und versteckt einen Knochen.

Der / Kein Mensch findet den Schatz / *Mond und vergräbt einen Hauptteil.

Der / Kein Verteidiger schießt den Elfmeter / *Versuch und trifft einen Torwart.

Der / Kein Opa pflückt den Apfel / *Rasenmäher und bereitet einen Obstsalat.

Der / Kein Mörder öffnet den Deckel / *Stuhl und versteckt einen Dolch.

Der / Kein Mechaniker erfindet den Toaster / *Husten und repariert einen Mixer.

Der / Kein Soldat gewinnt den Wettbewerb / *Schlaganfall und erhält einen Preis.

Der / Kein Naturschützer zerstört den Laster / *Satz und bekommt einen Verweis.

Der / Kein Lügner erzählt den Witz / *Rock und erntet einen Applaus.

Der / Kein Pfarrer putzt den Altar / *Dank und tauft einen Säugling.

Der / Kein Taucher nimmt den Schnorchel / *Kampf und sucht einen Taucheranzug.

Der / Kein Sekretär fährt den Bus / *Stock und hat einen Unfall.

Der / Kein Koch kauft den Salat / *Fühler und bereitet einen Hauptgang.

Der / Kein Fachmann programmiert den Computer / *Regenschirm und erhält einen Vertrag.

Der / Kein Schornsteinfeger reinigt den Kamin / *Roman und säubert einen Ofen.

Der / Kein Hausmeister repariert den Briefkasten / *Bauchnabel und fegt einen Boden.

Der / Kein Autor schreibt den Kommentar / *Spielplatz und veröffentlicht einen Roman.

Der / Kein Tenor singt den Song / *Bierdeckel und vergisst einen Text.

Der / Kein Putzmann kauft den Lappen / *Mord und reinigt einen Tisch.

Der / Kein Direktor trägt den Schal / *Asphalt und hat einen Rucksack.

Der / Kein Parkwächter beobachtet den Überfall / *Ausdruck und ruft einen Polizisten.

Der / Kein Anwalt verteidigt den Fall / *Warteraum und gewinnt einen Prozess.

Der / Kein Vorsitzende kritisiert den Beschluss / *Baum und erntet Beifall.

Der / Kein Raucher riskiert den Herzinfarkt / *Zeiger und raucht einen Zigarillo.

Der / Kein Moderator kommentiert den Vorfall / *Sack und zeigt einen Bericht.

Der / Kein Gärtner bewässert den Garten / *Planet und entfernt einen Schädling.

Der / Kein Verehrer spendiert den Opernbesuch / *Bahndamm und bekommt einen Kuss.

Der / Kein Bürgermeister organisiert den Ausflug / *Blitzschlag und erhält einen Preis.

5.2 Experiment 2: Double Nominative Violation

Der / Kein Mann verkauft den / *der Wagen und kauft einen Bus.

Der / Kein Bauer bestellt den / *der Acker und kauft einen Traktor.

Der / Kein Graf spendiert den / *der Pokal und erhält einen Orden.

Der / Kein Lehrer liest den / *der Aufsatz und vergibt einen Preis.

Der / Kein Maler wäscht den / *der Pinsel und mischt einen Farbton.

Der / Kein Tänzer übt den / *der Tango und verletzt einen Zeh.

Der / Kein Schüler kauft den / *der Kuchen und probiert einen Keks.

Der / Kein LKW-Fahrer singt den / *der Song und übersieht einen Bordstein.

Der / Kein Junge kauft den / *der Ball und beginnt einen Sportkurs.

Der / Kein Tiger beherrscht den / *der Dschungel und verteidigt einen Knochen.

Der / Kein Händler erfüllt den / *der Auftrag und erhält einen Bonus.

Der / Kein Schulleiter unterschreibt den / *der Bescheid und bewilligt einen Ausflug.

Der / Kein Sänger studiert den / *der Text und übt einen Auftritt.

Der / Kein Diener kocht den / *der Tee und bügelt einen Anzug.

Der / Kein Stürmer schießt den / *der Elfmeter und verletzt einen Fuß.

Der / Kein Student schreibt den / *der Aufsatz und findet einen Lösungsweg.

Der / Kein Fahrer rammt den / *der Pfeiler und streift einen Baum.

Der / Kein Angestellte beantwortet den / *der Brief und zahlt einen Aufpreis.

Der / Kein Chirurg unterstützt den / *der Eingriff und operiert einen Bypass.

Der / Kein Azubi isst den / *der Keks und trinkt einen Kaffee.

Der / Kein Chef bewilligt den / *der Antrag und diskutiert einen Entwurf.

Der / Kein Präsident verurteilt den / *der Anschlag und fasst einen Beschluss.

Der / Kein Lokführer stört den / *der Bahnverkehr und riskiert einen Rauswurf.

Der / Kein Kellner serviert den / *der Braten und bringt einen Wein.

Der / Kein Pförtner verliert den / *der Schlüssel und sucht einen Hinweis.

Der / Kein Mörder versteckt den / *der Leichnam und zerstört einen Beweis.

Der / Kein Schaffner kontrolliert den / *der Fahrausweis und erwischt einen Mann.

Der / Kein Clown sucht den / *der Lolli und findet einen Hut.

Der / Kein Zöllner kontrolliert den / *der Koffer und findet einen Dolch.

Der / Kein Angeklagter geht den / *der Rechtsweg und gewinnt einen Prozess.

Der / Kein Sportler umgeht den / *der Prozess und startet einen Wettkampf.

Der / Kein Gangster vergisst den / *der Handschuh und erschlägt einen Zeugen.

Der / Kein Lotse betritt den / *der Tower und bekommt einen Auftrag.

Der / Kein Bauherr bekommt den / *der Schlüssel und beginnt einen Umzug.

Der / Kein Informatiker repariert den / *der Computer und installiert einen Drucker.

Der / Kein Konditor säubert den / *der Ofen und reinigt einen Tisch.

Der / Kein Junge trinkt den / *der Schnaps und probiert einen Wein.

Der / Kein Kommissar rekonstruiert den / *der Mord und löst einen Fall.

Der / Kein Bote bringt den / *der Scheck und verlangt einen Beleg.

Der / Kein Hausmeister nimmt den / *der Besen und fegt einen Boden.

Der / Kein Hase frisst den / *der Salat und verschmäht einen Kohl.

Der / Kein Arzt behandelt den / *der Krebs und verschreibt einen Eingriff.

Der / Kein Tankwart misst den / *der Ölstand und wechselt einen Reifen

Der / Kein Richter sieht den / *der Stuhl und verurteilt einen Mörder

Der / Kein Dozent trinkt den / *der Wein und isst einen Kuchen.

5.3 Experiment 3: NPI Licensing[125]

Kein / *Der Bauer hat den Bären jemals / durchaus erschossen.

Kein / *Der Tänzer hat den Zuschauer jemals / durchaus beschimpft.

Kein / *Der Hund hat den Mann jemals / durchaus gebissen.

Kein / *Der Lehrer hat den Schüler jemals / durchaus gelobt.

Kein / *Der Torwart hat den Trainer jemals / durchaus getreten.

Kein / *Der Junge hat den Goldfisch jemals / durchaus gefüttert.

Kein / *Der Prinz hat den Frosch jemals / durchaus geküsst.

Kein / *Der Gärtner hat den Maulwurf jemals / durchaus gefunden.

Kein / *Der Rentner hat den Briefträger jemals / durchaus gesehen.

Kein / *Der Rennfahrer hat den Mechaniker jemals / durchaus gegrüßt.

Kein / *Der Präsident hat den Minister jemals / durchaus belogen.

Kein / *Der Mensch hat den Löwen jemals / durchaus besiegt.

Kein / *Der Affe hat den Wärter jemals / durchaus umarmt.

Kein / *Der Wähler hat den Politiker jemals / durchaus gemocht.

Kein / *Der Gorilla hat den Jäger jemals / durchaus verfolgt.

Kein / *Der Busfahrer hat den Radfahrer jemals / durchaus angefahren.

Kein / *Der Autor hat den Leser jemals / durchaus enttäuscht.

Kein / *Der Drache hat den Troll jemals / durchaus verehrt.

Kein / *Der Psychologe hat den Mann jemals / durchaus verstanden.

Kein / *Der Verkäufer hat den Kunden jemals / durchaus beraten.

Kein / *Der Arzt hat den Kranken jemals / durchaus behandelt.

Kein / *Der Detektiv hat den Mörder jemals / durchaus erwischt.

[125] *-indexed definite determiner indicates that the sentences including NPI "jemals" ('ever') are ill-formed. Sentences including the positive polarity item (PPI) "durchaus" ('especially') are ill-formed when the negative quantifier precedes the PPI.

Kein / *Der Hund hat den Einbrecher jemals / durchaus erschreckt.

Kein / *Der Fan hat den Star jemals / durchaus berührt.

Kein / *Der Zauberer hat den König jemals / durchaus verhext.

Kein / *Der Forscher hat den Kannibalen jemals / durchaus untersucht.

Kein / *Der Segler hat den Meeresgott jemals / durchaus gesehen.

Kein / *Der Hase hat den Igel jemals / durchaus überholt.

Kein / *Der Hausmeister hat den Mieter jemals / durchaus angeschrien.

Kein / *Der Wirt hat den Gast jemals / durchaus bedient.

Kein / *Der Freund hat den Müllmann jemals / durchaus angesprochen.

Kein / *Der Kommilitone hat den Dozenten jemals / durchaus belogen.

Kein / *Der Klient hat den Anwalt jemals / durchaus unterstützt.

Kein / *Der Drache hat den Ritter jemals / durchaus besiegt.

Kein / *Der Hund hat den Fuchs jemals / durchaus verjagt.

Kein / *Der Schauspieler hat den Regisseur jemals / durchaus belauscht.

Kein / *Der Mitarbeiter hat den Chef jemals / durchaus angelächelt.

Kein / *Der Fachmann hat den Laien jemals / durchaus überzeugt.

Kein / *Der Angeklagte/r hat den Richter jemals / durchaus bestochen.

Kein / *Der Eskimo hat den Eisbären jemals / durchaus gesehen.

Kein / *Der Fasan hat den Flamingo jemals / durchaus angegriffen.

Kein / *Der Pinguin hat den Seelöwen jemals / durchaus besucht.

Kein / *Der Frisör hat den Politiker jemals / durchaus bedient.

Kein / *Der Händler hat den Käufer jemals / durchaus angerufen.

Kein / *Der Soldat hat den Bürger jemals / durchaus angeschossen.

6 Appendix: Article

Eine EKP-Studie zum kindlichen Zweitspracherwerb

Juliane Domke

1. Einleitung

Die Frage nach dem Einfluss des Erwerbsalters auf den Zweitspracherwerb (im Folgenden L2-Erwerb) wird seit mehreren Dekaden diskutiert und es herrscht mittlerweise Konsens darüber, dass L2-Lerner[1] mit jüngerem Erwerbsalter gegenüber L2-Lernern mit spätem Erwerbsalter im Vorteil sind. Dieser Vorteil bezieht sich sowohl auf den Verlauf, als auch auf das Endergebnis des L2-Erwerbs. Nach dem Ansatz der von Lenneberg (1967) postulierten kritischen Periode[2] für den L2-Erwerb, wird deren Endpunkt mit dem Eintritt der Pubertät – einem durchschnittlichen Alter von zwölf Jahren – gleichgesetzt. Diesbezüglich wurde innerhalb der letzten ca. 30 Jahre ein robuster Einfluss des Alters auf den L2-Erwerb für verschiedene linguistische Bereiche empirisch untersucht und berichtet (siehe u.a. Hyltenstam/Abrahamson: 2003 für einen Überblick). Dabei hat sich herausgestellt, dass unterschiedliche kritische Perioden für unterschiedliche linguistische Domänen gelten (siehe u.a. Singleton: 2005), doch hält der von Lenneberg postulierte Endpunkt der (einen) kritischen Periode von zwölf Jahren den empirischen Ergebnissen nicht stand und wird allgemein auf ein Erwerbsalter im Bereich etwa bei sechs/sieben Jahre liegend herunterkorrigiert (u.a. Hyltenstam/Abrahamson: 2003, 575). Dieser sog. Alterseffekt bezieht sich vor allem auf das Ergebnis des L2-Erwerbs. Hier herrscht der Standpunkt vor, dass einem Lerner bis zu einem Erwerbsalter von ca. sechs Jahren Muttersprachlichkeit in der Zweitsprache (im Folgenden L2) garantiert ist, danach nicht mehr. Welche Prozesse zu welchen Zeitpunkten während dieser kritischen Periode(n) betroffen sind, ist weniger transparent. In jedem Fall bezeichnet das L2-Erwerbsalter um ca. sechs Jahre einen kritischen Punkt der zwischen frühem (< sechs Jahre) und spätem L2-Erwerb (> sechs Jahre) unterscheiden lässt.[3]

In den letzten Jahren konzentriert sich das Forschungsinteresse auf einen Bereich, der traditionell gesehen in den frühen L2-Erwerb fällt, nämlich auf den kindlichen L2-Erwerb (Meisel: 2008; 2010; Rothweiler: 2007). Man spricht vom kindlichen L2-Erwerb, wenn ein Kind eine (Zweit-)Sprache nicht simultan – also zwei (oder mehr) Muttersprachen gleichzeitig – erwirbt, sondern wenn die L2 später in der Kindheit erlernt wird, ab einem Alter von ca. vier Jahren. Die Idee, dass der frühe L2-Erwerb sich in

---

[1] Ich verwende im vorliegenden Artikel die männliche Form von Lerner die als generische, die sich sowohl auf männliche als auch weibliche Lerner bezieht.

[2] In der Literatur findet man synonym die Termini kritische Periode, kritische Phase und auch sensitive Phase.

[3] Die Unterscheidung zwischen frühem und spätem L2-Erwerb ist in der Literatur nicht einheitlich (siehe u.a. Meisel, 2010, der von spätem L2-Erwerb ab einem Erwerbsalter bei ca. zehn Jahren liegend spricht). In diesem Artikel beziehe ich mich auf die Bezeichnung des späten L2-Erwerbs ab einem Erwerbsalter bei ca. sechs/sieben Jahren liegend (siehe u.a. Hyltenstam/Abrahamson, 2003).

die Phasen simultan (ca. null bis drei Jahre) und sukzessive bzw. kindlich (ca. vier bis sechs Jahre) aufteilt, basiert auf der Annahme, dass zentrale Spracherwerbsprozesse bezüglich der Muttersprache bereits in den ersten zwei bis drei Lebensjahren abschließen (u.a. Meisel: 2010; Rothweiler: 2007). Dementsprechend besteht somit die Annahme, dass die Phase des kindlichen L2-Erwerbs innerhalb der Diskussion um die kritische Periode eine Schlüsselstellung einnimmt und zwar dahingehend, dass die traditionelle Unterscheidung zwischen frühem und spätem L2-Erwerb weiterhin bestehen bleibt, die Phase des frühen L2-Erwerbs jedoch modifiziert werden muss. Aufgrund der erst jungen Forschungsgeschichte dieses Bereichs, gibt es derzeit noch wenig Erkenntnis darüber, inwieweit der kindliche L2-Erwerb Gemeinsamkeiten und Unterschiede jeweils zum simultanen und späten L2-Erwerb aufweist und welche Eigenschaften die Phase des kindlichen L2-Erwerbs als solche ausmachen bzw. ihr eigen sind.

Das vorliegende Kapitel beschäftigt sich diesbezüglich mit der Frage ob der kindliche L2-Erwerb Unterschiede zum simultanen und/oder späten L2-Erwerb aufzeigt. Dazu wurde mittels der EKP-Methode untersucht, ob sich Unterschiede in der Sprachverarbeitung zeigen, die auf den Einfluss des Erwerbsalters zurückzuführen sind, wenn die L2 zu unterschiedlichen Zeitpunkten während der Kindheit erworben wurde. Der Artikel baut sich folgendermaßen auf: Zunächst wird ein kurzer Überblick über die Annahmen zur Phase des kindlichen L2-Erwerbs gegeben. Dem folgt ein kurzer Überblick über bisherige relevante EKP-Ergebnisse in Bezug auf den L2-Erwerb und die dann folgend vorgestellte Studie.

2. Vorbemerkungen

2.1 Zum kindlichen L2-Erwerb

Man weiß, dass sich der späte L2-Erwerb (> sechs Jahre) sowohl im Verlauf als auch im Ergebnis vom simultanen L2-Erwerb unterscheidet. Meisel (2008; 2010) nimmt an, dass sich diese Unterschiede bereits in frühen Stadien des L2-Erwerbs zeigen und dass der kindliche L2-Erwerb in vielerlei Hinsicht mehr Gemeinsamkeiten zum späten L2-Erwerb aufzeigt, als zum simultanen L2-Erwerb. Erklärungen für diese Unterschiede beziehen sich im Wesentlichen auf zwei Grundannahmen: Zum einen wird angenommen, dass es einschlägige kognitive Entwicklungsstufen innerhalb dieser Phase gibt, die zu Unterschieden z.B. in der Lernstrategie und in dem was ein Lerner mitbringt führen (Meisel: 2008; Rothweiler: 2007). Z.B. unterscheidet sich ein Kind von vier Jahren markant von einem sechs Monate alten Säugling bezüglich der Voraussetzungen für den L2-Erwerb. Das vierjährige Kind hat neben umfassenden Alltagserfahrungen die Grundzüge einer (Mutter-)Sprache bereits erworben und verfügt über ausgeprägte kognitive Fähigkeiten wie z.B. spezifische Lernmechanismen. Zum anderen besteht die Vermutung, dass sich die sog. Spracherwerbsmechanismen schon im frühen Kindheitsalter dahingehend verändern, dass sie mit ansteigendem Erwerbsalter sukzessive abnehmen (vgl. Rothweiler: 2007). Diese Annahme beruht auf der Evidenz der hohen neuronalen Plastizität, die mit fortschreitender zerebraler Lateralisierung immer weiter abbaut. Einem Kind zwischen Geburt und drei Jahren stehen daher aufgrund hoher neuronaler Flexibilität multiple Fähigkeiten zum Spracherwerb zur Verfügung. In einem Alter zwischen vier und sechs Jahren verliert die neuronale Plastizität kontinuierlich an Flexibilität, was mit einem sukzessiven Abbau der Spracherwerbsmechanismen einhergeht. Ab einem Alter von sieben Jahren erfolgt dieser Abbau rapider, spätestens mit zehn Jahren schließt sich die sog. kritische Phase für den L2-Erwerb, die ein Ergebnis der muttersprachlichen Kompetenz oder Sprachbeherrschung garantiert (vgl. Rothweiler: 2007; Meisel: 2010).

Neben den Unterschieden bezüglich der Spracherwerbsmechanismen, besteht auch die Annahme über eine Reihe nichtsprachlicher Einflüsse, die zu Unterschieden zwischen simultanem, kindlichen und späten L2-Erwerb führen. Während der unauffällige (simultane) L2-Erwerb weitestgehend unabhängig von nichtsprachlichen Einflussfaktoren verläuft, ist dies beim sukzessiven (vor allem beim späten) L2-Erwerb nicht der Fall (vgl. u.a. Long: 1990; Rothweiler: 2007). Solche Einflussfaktoren beziehen sich neben dem Erwerbsalter u.a. auch auf das soziale Umfeld (gemeint ist hier besonders die Menge und Qualität des L2 Inputs), die Motivation des Lerners und den Stand der Sprachbeherrschung (Rothweiler: 2007, 122). Das bedeutet, dass mit steigendem Erwerbsalter die Rolle der nichtsprachlichen Faktoren zunimmt und sich dementsprechend auch die individuelle Lernervariation erhöht (Long: 1990). Da nicht klar ist, welchen Umfang die nichtsprachlichen Einflussfaktoren bezüglich des kindlichen L2-Erwerbs einnehmen, soll dieser Punkt im Rahmen des vorliegenden Artikels nicht weiter diskutiert werden, dennoch sei festgehalten, dass diesbezüglich weiterer Untersuchungsbedarf besteht.

Empirische Evidenz zum kindlichen L2-Erwerb ist wie eingangs erwähnt derzeit noch wenig vorhanden und die Ergebnisse sind nicht einheitlich. Gründe dafür sind u.a., dass dieser Forschungsgegenstand ein noch junger ist, dass die kritische Periode in unterschiedlichen linguistischen Domänen unterschiedliche Zeitfenster bedient und vor allem dass die Untersuchungsmethoden sehr unterschiedlich sind (für einen Überblick siehe Meisel: 2010).

2.2 Zur Untersuchungsmethode

Hier soll kurz die EKP-Methode vorgestellt werden. Diese Methode beruht auf der Evozierung von Unterschieden in der Sprachverarbeitung, die z.B. auf den potenziellen Einfluss des Erwerbsalters zurückzuführen sein können. Es handelt sich bei dieser Untersuchungsmethode zu (L2-)Spracherwerbsmechanismen um sog. Quasiexperimente, „d.h. Versuchspersonen können nicht zufällig den experimentellen Bedingungen zugeordnet werden" (Müller/Hirotani/Friederici: 2006, 180). Doch hat sich die EKP-Methode zur Evozierung von linguistischen Verarbeitungsunterschieden auch bei L2-Lernern mittlerweile als nützliches Verfahren etabliert, denn, „[b]ringt man solche Ergebnisse in Zusammenhang mit gut kontrollierten Probandenvariablen, wie z.B. Erwerbsalter […], so können gezielt bestimmte Zusammenhangsvermutungen getestet werden" (Müller/Hirotani/Friederici: 2006, 180).

Die Sprachverarbeitung z.B. eines Satzes ist ein kontinuierlicher Prozess, der in zeitliche Bereiche eingeteilt werden kann. Dabei unterscheidet man zwischen frühen und späten sowie syntaktischen und semantischen Verarbeitungsprozessen (siehe u.a. Friederici: 1995; 2002). Das erwartete Muster, das z.B. bei der Verarbeitung morphosyntaktischer Anomalien innerhalb eines Verletzungsparadigmas[4] evoziert würde, ist das sog. LAN-P600 Muster. Die LAN (links anteriore Negativierung) ist eine Komponente, die relativ früh in der Verarbeitung zwischen ca. 300 und 500 ms nach Präsentation des

[4] Man spricht von einem Verletzungsparadigma, wenn EKPs syntaktisch korrekter Konstruktionen mit ihren analog syntaktisch inkorrekten verglichen werden.

kritischen Elements auftritt. Sie entspricht automatischen Prozessen der Strukturregelverarbeitung und wird bei morphosyntaktischen Verletzungen beobachtet (u.a. Friederici: 2002). Oft folgt dem durch eine morphosyntaktische Verletzung ausgelösten Korrelat LAN die sog. P600-Komponente (Osterhout/Holcomb: 1992). Sie entspricht einer Positivierung, die zwischen ca. 600 und 800 ms nach Präsentation des kritischen Elements auftritt und wird mit Reparatur- und Reanalysemechanismen (Friederici: 1995) oder auch mit erhöhter Verarbeitungskomplexität (u.a. Friederici/Hahne/Saddy: 2002) sowie mit Schwierigkeiten der syntaktischen Integration (Kaan/Harris/Gibson/Holcomb: 2000) assoziiert. Zusammenfassend: Die Verarbeitung morphosyntaktischer Strukturen/ Strukturverletzungen findet in zwei Schritten statt: zuerst wird die fehlerhafte morphosyntaktische Information detektiert (LAN), dies passiert weitestgehend automatisiert. Im zweiten Schritt wird dann anhand hoch kontrollierter Sprachverarbeitungsprozesse versucht, die fehlerhafte Struktur zu reparieren (P600).

EKP-Studien zur morphosyntaktischen Verarbeitung bei L2-Lernern berichten Unterschiede zwischen frühem und späten L2-Erwerb (für einen Überblick siehe Müller/Rüschemeyer/Friederici: 2006). Dabei wird gezeigt, dass die größten Unterschiede die frühen Verarbeitungsprozesse betreffen. Als Grund dafür wird angenommen, dass L2-Lerner einen höheren Kapazitätenanspruch an die Sprachverarbeitungsprozesse stellen, die mittels EKP sichtbar werden. Z.B. berichtet Hahne (2001), dass besonders bei frühen syntaktischen Effekten (wie z.B. der LAN) dahingehend Unterschiede auftreten, dass L2-Lerner mit spätem Erwerbsalter diese frühen Effekte nicht zeigen. Die Abwesenheit dieser frühen syntaktischen Effekte erklärt sie damit, dass Prozesse zum Aufbau einer Struktur, die zur Detektion solcher Fehler führen, bei Lernern mit spätem Erwerbsalter nicht durchlaufen werden, bzw. diese (noch) nicht automatisiert sind. Einen alternativen Vorschlag für die Abwesenheit dieser Effekte machen Osterhout, McLaughlin, Pitkanen, Frenck-Mestre und Molinaro (2006), die annehmen, dass L2-Lerner eine höhere individuelle Lernervariation innerhalb der EKPs zeigen und durch die Mittelung zum Grand Average[5] diese Effekte, die einzelne L2-Lerner durchaus evozieren, verschwinden. Der Einfluss des Erwerbsalters auf die morphosyntaktische Verarbeitung basiert also auf der Vermutung, dass wenn Lerner mit spätem Erwerbsalter morphosyntaktische Verletzungen detektieren, die Effekte kleiner sind, da ihre Verarbeitung höhere Ansprüche an den Prozessor stellt und/oder das Korrelat aufgrund erhöhter Lernervariation geschwächt wird.

Dagegen zeigt sich der Einfluss des Erwerbsalters bei späten Verarbeitungskomponenten wie der P600 nicht so stark. Jedoch scheint es in diesem Fall eine Korrelation mit dem Kenntnisstand der jeweiligen L2 zu geben. So wird u.a. berichtet, dass die P600 stabil auftritt, wenn der Kenntnisstand einer L2 entsprechend hoch ist (Hahne: 2001; Müller/Hirotani/Friederici: 2006). Abweichungen zur muttersprachlichen Verarbeitung der P600 zeigen sich meist in einer zeitlichen Verzögerung des Effektes. Es ist also anzunehmen, dass ein L2-Lerner mit einem hohen Kenntnisstand in der L2 auf kontrollierte Sprachverarbeitungsprozesse wie z.B. die Reparatur eines morphosyntaktischen Fehlers zurückgreifen kann. Zeitliche Verzögerungen dieser Prozesse im Gegensatz zur muttersprachlichen Verarbeitung können auch hier auf den erhöhten Kapazitätenanspruch an den Prozessor zurückgeführt werden.

[5] Der Grand Average (Gesamtdurchschnittswert) ergibt sich aus der Mittelung der Durchschnittswerte der jeweiligen Bedingungen erst pro Versuchsperson, dann über alle Versuchspersonen.

EKP-Studien, die den Einfluss des Erwerbsalters auf die Sprachverarbeitung in Bezug auf den Bereich des kindlichen L2-Erwerbs untersuchen, sind mir nicht bekannt. Weber-Fox und Neville (1996) ist die mir bisher einzig bekannte Studie, in dem EKP-Ergebnisse einer Lernergruppe (Chinesisch (L1) – Englisch (L2)) mit Erwerbsalter von vier bis sechs Jahren berichtet werden. Die Autorinnen untersuchten jedoch nicht explizit diesen Bereich, sondern allgemein den Einfluss des Erwerbsalters auf die Verarbeitung unterschiedlicher sprachlicher Strukturen. Sie finden u.a. dass sich der Einfluss des Erwerbsalters bei der Phrasenstrukturverarbeitung bereits früh ab einem Alter zwischen vier und sechs Jahren im L2-Erwerb manifestiert. Sie zeigen, dass die Ergebnisse der Gruppe mit einem Erwerbsalter von vier bis sechs Jahren („Gruppe 4-6") Unterschiede zu der Gruppe mit einem Erwerbsalter von 1-3 Jahren („Gruppe 1-3") sowohl in den behavioralen als auch in den elektrophysiologischen Daten aufweisen. Bezüglich der behavioralen Daten berichten sie für die „Gruppe 4-6" schlechtere Ergebnisse in den Akkuratheitsdaten, was sie als Trend in Bezug auf einen geringeren Kenntnisstand in der L2 interpretieren. Weiterhin geben sie an, dass die Akkuratheitsdaten der „Gruppe 4-6" Gemeinsamkeiten zu den Lernergruppen mit spätem Erwerbsalter („Gruppe 7-10", „Gruppe 11-13") aufweisen. Sie konstatieren weiter, dass sich auch bezüglich früher EKP-Korrelate Unterschiede in der Verarbeitung von morphosyntaktischer Information schon ab einem Erwerbsalter von 4 Jahren feststellen lassen, in dem frühe EKP-Komponenten entweder gar nicht auftraten bzw. wenn sie auftraten, eine unterschiedliche Verteilung aufwiesen. Späte EKP-Komponenten in den Lernergruppen „Gruppe 4-6" und „Gruppe 7-10" wiesen keine Unterschiede zu denen von Muttersprachlern auf, ein signifikanter Unterschied wurde hier (erst) für Lernergruppen mit einem Erwerbsalter > elf Jahre gefunden.

Im verbleibenden Teil dieses Artikels soll eine Studie vorgestellt werden, in der die Sprachverarbeitung einer morphosyntaktischen Strukturverletzung von L2-Lernern mit unterschiedlichem Erwerbsalter (simultaner, kindlicher und später L2-Erwerb) untersucht wurde.

3. Studie

3.1 Untersuchungsgegenstand, Problemstellung und Vorhersagen

In dieser Studie mit polnisch-deutschen Lernern wurde mittels EKP-Methode untersucht, ob sich aufgrund unterschiedlichen Erwerbsalters Unterschiede bei der L2-Verarbeitung zeigen, wenn der Kasus am direkten Akkusativobjekt nicht adäquat bzw. fehlerhaft markiert ist. Es ging darum, herauszufinden, ob sich unterschiedliche Verarbeitungsmuster zeigen, wenn die L2 simultan, später im Kindheitsalter oder spät, also nach einem Alter von sechs Jahren, erworben wurde.

Allgemein geht man davon aus, dass die Kasusfunktionen für strukturellen Kasus früh im Erstspracherwerb – ab einem Alter von ungefähr zwei Jahren – operativ (Babyonyshev: 1993; Eisenbeiss/Bartke/Clahsen: 2006) bzw. vollständig erworben sind (Radford: 1990; für einen Überblick zum Kasuserwerb siehe Eisenbeiss/Narashiman/Voeikova: 2009). Dementsprechend dient Kasusmarkierung als solider Untersuchungsgegenstand, um mögliche Unterschiede zwischen simultanem und kindlichem (sowie spätem) L2-Erwerb festzustellen. Kinder, die von frühester Kindheit (Geburt bis drei

Jahre) an bilingual aufwachsen, erwerben unterschiedliche Systeme zur Kasusmarkierung simultan. Hier sollten die Grundlagen der Systeme der jeweiligen Sprachen mit einem Alter ab spätestens vier Jahren erworben sein. Nach der These von Meisel (2008; 2010), entspricht der kindliche L2-Erwerb eines weiteren Systems zur Kasusmarkierung dann nicht mehr dem des muttersprachlichen bzw. simultanen Erwerbs. Folglich sollten sich bezüglich der Detektion von Fehlern in der Kasusmarkierung Unterschiede bei deren Verarbeitung zeigen.

Das Deutsche markiert strukturellen Kasus am Nomen[6] meist durch einen adäquaten Artikel. Es ist manchmal der Fall, dass ein Nomen im Akkusativ eine Flexionsendung erhält (der Löwe$_{\text{Nominativ}}$ – den Löwen$_{\text{Akkusativ}}$), in der Regel wird aber der Akkusativ nur durch Artikelflexion angezeigt (der Richter$_{\text{Nominativ}}$ – den Richter$_{\text{Akkusativ}}$) (vgl. Wegener: 1995). Im Polnischen gibt es kein Artikelsystem zur näheren Bestimmung der Eigenschaften (z.B. Kasus) des Nomens. Hier wird struktureller Kasus (z.B. Akkusativ) direkt am Nomen markiert. Dementsprechend unterscheiden sich die beiden Sprachen in ihrer Markierung von strukturellem Akkusativ. Für den L2-Erwerb bedeutet das, dass polnische Lerner des Deutschen ein neues System zur Markierung von Akkusativ lernen müssen, nämlich, dass (maskuline) Nomen für Akkusativ nicht flektiert werden, aber einen anderen Artikel benötigen, als den des Nominativs. Entsprechend der Annahmen des Einflusses des Erwerbsalters auf den L2-Erwerb sollten sich Unterschiede in der Verarbeitung syntaktischer Strukturen, die eine fehlerhafte Kasusmarkierung beinhalten, zwischen Lernergruppen mit unterschiedlichem Erwerbsalter zeigen. Bezüglich dieser Untersuchung wird für den simultanen L2-Erwerb ein Verarbeitungsmuster LAN-P600 erwartet. Für den kindlichen und späten L2-Erwerb wird eine P600-Komponente, jedoch keine stabile LAN-Komponente erwartet. Die Vorhersagen beruhen auf der Annahme, dass wenn Kasusmarkierungsprozesse bereits mit zwei Jahren operativ und spätestens mit vier Jahren vollständig erworben sind, beim späteren Erwerb (≥ vier Jahre) eines weiteren Systems zur Kasusmarkierung höhere Ansprüche an die Verarbeitungskapazität gestellt werden. Diese höheren Ansprüche an die Verarbeitungskapazität sollten sich vor allem in der frühen Verarbeitungsphase zeigen, da Prozesse wie diese eine hohe Automatizität der Sprachverarbeitung verlangen. Späte Verarbeitungsprozesse, sollten von dem Einfluss des Erwerbsalters nicht so stark betroffen sein, hier werden keine bzw. nur geringfügige Unterschiede erwartet (siehe auch 2.2).

3.2 Versuchspersonen

30 polnische Muttersprachler (zehn davon männlich) mit Deutsch als L2 wurden getestet. Die Versuchspersonen (im Folgenden VP(n)) wurden bezüglich ihres Erwerbsalters in vier Gruppen aufgeteilt. Die Aufteilung wurde gemäß der (spekulativen) Erwerbsaltersbereiche Weber-Fox und Neville (1996) folgend unternommen: simultaner L2-Erwerb beinhaltete die VPn, die Deutsch in einem Alter von Geburt an bis hin zu drei Jahren erworben haben. Eine weitere Gruppe bilden die VPn mit einem Erwerbsalter zwischen vier und sechs Jahren. Diese entspricht dem kindlichen Zweitspracherwerb. VPn mit spätem Erwerbsalter (> 6 Jahre) teilen sich in zwei Gruppen, 7-10 Jahre und 11-13 Jahre (siehe auch Meisel: 2010). Alle VPn hatten einen Kenntnisstand der deutschen Sprache im C-Bereich gemäß dem gemeinsamen europäischen Referenzrahmen für

[6] Im vorliegenden Artikel beziehe ich mich nur auf maskuline Nomen.

Sprache (81-100 %), siehe auch Tabelle 1. Alle VPn waren zum Zeitpunkt der Erhebung zwischen 20 und 31 Jahre alt (Durchschnittsalter 24,8). Sie alle studierten an einer Berliner oder Brandenburger Universität, waren rechtshändig und hatten normale oder zu-normal-korrigierte Sicht.

Gruppe	**mittleres Alter**	**mittleres Erwerbsalter**	**mittlerer Kenntnisstand in %**
0-3 (n = 8)	24,53	1,3	97
4-6 (n = 7)	24,43	4,9	97
7-10 (n = 7)	26,5	9,1	94
11-13 (n = 8)	24,0	11,7	90

Tabelle 1: Zusammenfassung der experimentellen Gruppen

3.3 Stimulusmaterial

Das Stimulusmaterial bestand aus 45 Sätzen mit jeweils einer Länge von neun Worten. Jeder Satz kam zweimal vor, einmal in der korrekten Version (adäquate Kasusmarkierung am direkten Akkusativobjekt), einmal in der inkorrekten Version (fehlerhafte Kasusmarkierung), siehe auch (1). Zudem wurde die gleiche Anzahl (2x45) an sog. Distraktorsätzen präsentiert. Die Stimuli und Distraktoren wurden vor der Präsentation pseudo-randomisiert. Insgesamt hat jede VP 180 Sätze gelesen. Für die folgende Analyse (4.1-4.2) ergibt sich für die Auswertung eine Gesamtanzahl von 90 Sätzen pro VP, bzw. 45 Sätze pro Bedingung pro VP.

(1a) Der Autor schreibt den Roman und erhält einen Preis.
(1b) *Der Autor schreibt der Roman und erhält einen Preis.

Um Kasusambiguitäten zu kontrollieren standen alle Nominalphrasen im Singular und waren maskulin. Es wurde darauf geachtet, dass sich am kritischen Element (Akkusativobjekt) die Kasusmorphologie für Nominativ und Akkusativ nicht unterschied (siehe 3.1). Beide Sprachen, das Polnische und das Deutsche, verfügen über eine relativ freie Wortstellung, jedoch gilt in Hauptsätzen für die unmarkierte Wortstellung die SVO Abfolge bzw. im Deutschen die V2-Stellung. Hier gibt es oberflächenstrukturell keine sprachlichen Unterschiede, die sich in der Verarbeitung zeigen könnten. Dieses ermöglicht, dass eine negative Konfundierung mit dem eigentlichen Untersuchungsgegenstand weitgehend kontrolliert werden kann.

3.4 EEG Ableitung und Durchführung

Die Aufnahme des EEGs erfolgte durch Ag-AgCl-Elektroden in einer Frequenz von 250 Hz mit Impedanzen < 5 kΩ. Es wurde monopolar referenziell zum linken Mastoid (M1) von den folgenden Positionen nach dem standardisierten 10-20 System für Elektrodenpositionen (Sharbrough/Chatrian/Lesser/Lüders/Nuwer/Picton: 1991) abgeleitet: F7, F3, FZ, F4, F8, FC5, FCZ, FC6, T7, C3, C4, T8, CP5, CPZ, CP6, P7, P3, PZ, P4, P8, PO3, POZ, PO4 und OZ. Für das Elektrookulogram (EOG) wurde horizontal je eine Elektrode am linken/rechten äußeren Cantus und vertikal je eine Elektrode über/unter dem rechten Auge platziert. Das Experiment verteilte sich auf zwei Termine, die immer an unterschiedlichen Tagen stattfanden. Vorab der ersten Sitzung füllte jede VP einen Händigkeitsbogen aus (Oldfield: 1971). Zum ersten Termin wurde per C-Test[7] der Kenntnisstand der deutschen Sprache erhoben und eine Laborbelehrung durchgeführt. Zu Beginn der zweiten Sitzung füllte jede VP eine Einverständniserklärung aus, danach wurde das Elektrookulogramm, dann die der Kopfgröße der VP angepasste EEG-Kappe (Easycap GmbH: Hersching, Deutschland) vorbereitet. Während der EEG-Aufzeichnung saß die VP in einem abgedunkelten Raum vor einem Monitor (19 Zoll LCD-Monitor, Typ: Acer AL 1923), an dem die Stimuli (Wort für Wort) präsentiert wurden. Die VPn wurden dahingehend instruiert, die Sätze einzig nach deren grammatischen Wohlgeformtheit zu beurteilen. Um Händigkeitseffekte zu vermeiden, wurden die Antworttasten (korrekt/inkorrekt) alternierend pro VP pro Gruppe jeweils gewechselt – die Hälfte der VPn wählten mit der rechten Hand „korrekt", die andere Hälfte mit der linken Hand. Zunächst wurden 16 Übungssätze präsentiert, in denen die VP mit Hilfe der Versuchsleiterin daraufhin trainiert wurde, Bewegungen und Augenzwinkern auf den zeitlichen Bereich nach der Bewertung des aktuellen Satzes und vor der Präsentation des nächsten Satzes zu reduzieren. Das eigentliche Experiment unterteilte sich in fünf kurze Blöcke, die von Pausen, deren Länge die Versuchsperson selbst wählen konnte, unterbrochen wurden. Gemessen wurden die Akkuratheit und Reaktionszeiten der Antwortgaben, sowie die elektrophysiologischen Potenziale.

3.5 Datenanalysen

Behaviorale (Akkuratheitsraten und Reaktionszeiten der Antworten) und EKP-Daten wurden mit Hilfe des Softwarepaketes R 2.13.0 (R Development Core Team: 2011) ausgewertet. Für die Berechnung der behavioralen Daten wurde ein lineares gemischtes Modell (LMM = „linear mixed model") mit der Funktion lmer (Paket lme4, siehe Bates/Maechler/Bolker: 2011) gewählt. LMMs haben gegenüber der traditionellen Varianzanalyse (ANOVA) u.a. die Vorteile, dass sie zufällige Faktoren wie Versuchsperson und Stimuluselement spezifizieren, nicht auf die Analyse von Faktoren mit fester Anzahl von kategorischen Stufen beschränkt sind, sondern Tests bezüglich kontinuierlicher Variablen (Kovariate) erlauben und weniger anfällig gegenüber Datenverlust aufgrund experimenteller Manipulation sind (siehe Kliegl/Wei/Dambacher/Yan/Zhou: 2011, 2; Baayen: 2008, 241f). In die Berechnung der behavioralen Daten gingen Bedingung und Gruppe als feste Faktoren und Versuchsperson sowie Stimuluselement als zufällige Faktoren ein. Die abhängigen Faktoren waren Akkuratheit bzw. Reaktionszeit. Da die Verteilung der Reaktionszeiten keiner Normalverteilung folgte, wurde mit logarithmierten Zeiten der Rohdaten gerechnet.

Für die Berechnung der EKP-Daten wurde eine mehrfaktorielle ANOVA gewählt (wiederholte Messungen). Gemittelt wurde erst pro Bedingung für jede VP und dann pro Bedingung über alle VPn (Grand Average). Gemessen wurde die Abhängigkeit des elektrophysiologischen Potenzials von den festen Faktoren Bedingung (Bedingung), Gruppe (Gruppe) und „Region of Interest" (ROI)[8].

---

[7] Der C-Test wurde dem standardisierten C-Test Korpus des Sprachenzentrums der Humboldt-Universität zu Berlin entnommen. Er bestand aus fünf kurzen Lückentexten mit je 20 Lücken, die unter Zeitvorgabe per Hand ausgefüllt wurden.

[8] ROI (Region of Interest): Elektroden, die nahe bei einander liegen, werden zu einer ROI zusammengefasst. Dabei unternimmt man Kreuzungen von Hemisphäre (rechts / links) und Position (vorne / hinten). Für die vorliegende Auswertung wurden folgende ROIs festgelegt: Mittellinie

4. Ergebnisse

4.1 Behaviorale Daten

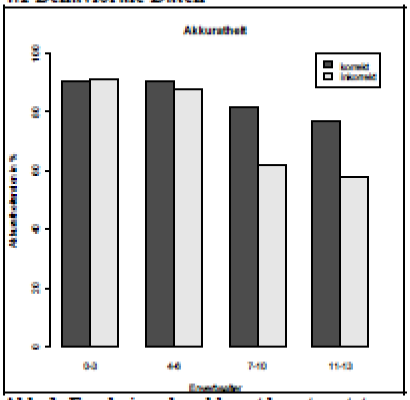


Abb. 1: Ergebnisse der akkurat beantworteten Bedingungen für alle vier experimentellen Gruppen

Gruppe / Erwerbsalter	Akkuratheit der Bedingungen in %	
	korrekt	inkorrekt
0-3	90,42	91,23
4-6	90,64	87,68
7-10	81,86	61,88
11-13	76,77	57,83

Tabelle 2: mittlere Akkuratheiten (in Prozent)

Abb. 1 illustriert die Unterschiede in den Akkuratheiten zwischen den Gruppen. Die statistische Analyse zeigt weder einen signifikanten Haupteffekt für Bedingung noch für Gruppe. Gruppe 11-13 zeigte einen Trend bezüglich eines Gruppenunterschieds (*Koeff*= -0,1287, *StF*=0,073, *t*=-1,76, *pMCMC*=0,06). Diese Gruppe zeigt die geringsten Akkuratheitswerte für beide Bedingungen (siehe Tabelle 2). Weiterhin zeigten sich signifikante Bedingungsunterschiede für die Gruppe 7-10 (*Koeff*=-0,0208, *StF*=0,0472, *t*=-4,41, *pMCMC*=0,0001) sowie für die Gruppe 11-13 (*Koeff*=-0,0211, *StF*=0,0448, *t*=-4,72, *pMCMC*=0,0001). In diesen beiden Gruppen waren die Akkuratheitsraten für die inkorrekte Bedingung signifikant schlechter als für die korrekte Bedingung. Die Gruppen 0-3 und 4-6 unterschieden sich nicht in den Akkuratheitsdaten.

---

mit FZ, CZ und PZ; anterior links mit F3, C3, FC5; anterior rechts mit F4, C4 und FC6; posterior links mit P3, PO3, CP5; sowie posterior rechts mit P4, PO4 und CP6.

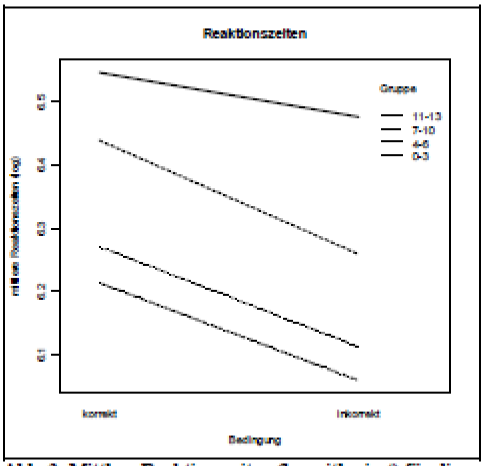


Abb. 2: Mittlere Reaktionszeiten (logarithmiert) für die experimentellen Bedingungen abgetragen für alle vier Gruppen

Abb. 2 zeigt die mittleren Reaktionszeiten der experimentellen Bedingungen für die vier experimentellen Gruppen. Die statistische Analyse zeigt einen signifikanten Haupteffekt für Bedingung (*Koeff*=-0,1487, *StF*=0,0467, *t*=-3,18, *pMCMC*=0,0012). Die Reaktionszeiten der inkorrekten Bedingung waren in allen Gruppen schneller, als die der korrekten Bedingung. Bezüglich eines Gruppeneffektes zeigt die Gruppe 7-10 einen Trend auf (*Koeff*=0,2074, *StF*=0,1232, *t*=1,68, *pMCMC*=0,07), für die Gruppe 11-13 ist dieser Unterschied signifikant (*Koeff*=0,3078, *StF*=0,1158, *t*=2,66, *pMCMC*=0,006). Weiterhin zeigt die Gruppe 11-13 einen Trend für einen Unterschied zwischen den Bedingungen (*Koeff*=0,1047, *StF*=0,0613, *t*=1,71, *pMCMC*=0,08). Diese Gruppe unterscheidet sich von den anderen Gruppen dahingehend, dass sie sowohl die langsamsten Reaktionszeiten als auch den geringsten zeitlichen Unterschied in der Schnelligkeit der Beantwortung zwischen den Bedingungen zeigt. Unterschiede zwischen den beiden frühen Lernergruppen 0-3 und 4-6 wurden auch hier nicht gefunden.

4.2 EKP-Daten

Zur bildlichen Analyse der EKP-Daten wurde die Auswertungssoftware Brain Vision Analyser (Brain Products) eingesetzt. Die anschließende statistische Berechnung fand mit R.2.13 statt (siehe 3.5). In die Analyse ging die Berechnung der Abhängigkeit des elektrophysiologischen Potenzials von den festen Faktoren Bedingung, Gruppe und ROI ein. Gemäß der Vorhersagen (siehe 3.1) und der visuellen Inspektion wurden zwei Zeitfenster geschnitten: 400-500 ms (LAN) und 600-800 ms (P600). Abb. 3 zeigt die Grand Average EKPs für die jeweiligen experimentellen Gruppen.

Appendix 6

Eine EKP-Studie zum kindlichen Zweitspracherwerb 197

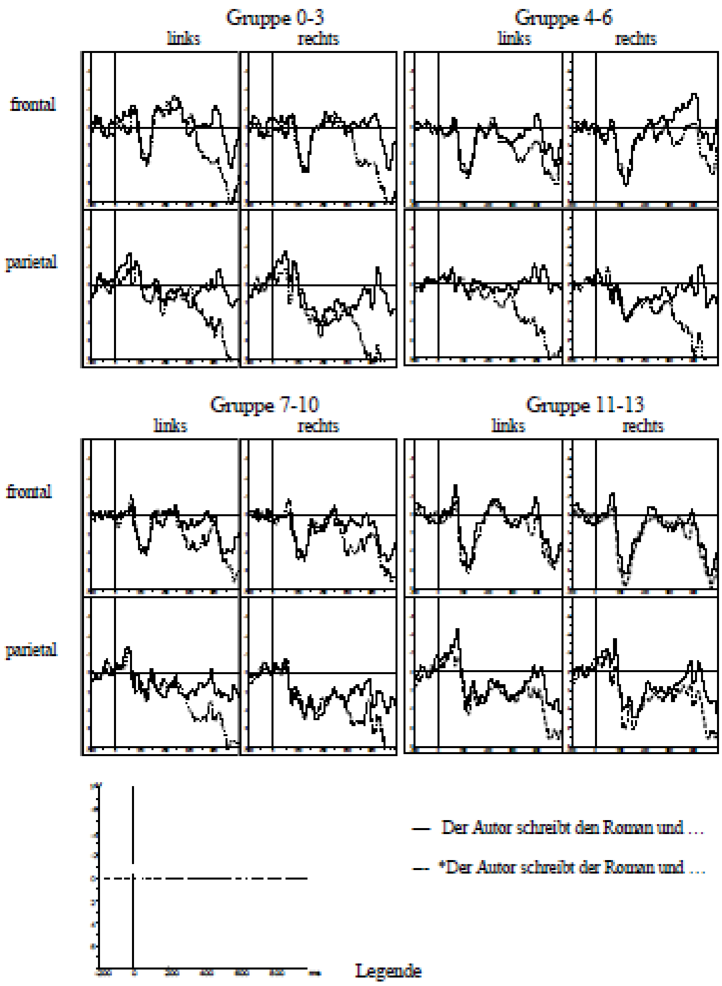


Abb. 3: Grand Average EKPs abgetragen jeweils für die vier experimentellen Gruppen; -200 ms ab dem kritischen Element (kursiv). Die gestrichelte Linie illustriert die evozierte Reaktion auf die fehlerhafte Kasusmarkierung.

LAN: Folgend der visuellen Inspektion (siehe Abb. 3) wurde für die Berechnung der LAN das Zeitfenster 400-500 ms gewählt. Die globale Analyse in diesem Zeitfenster zeigt signifikante Haupteffekte für Bedingung ($F(1,829)=11,57$; $p<0,001$) und ROI ($F(4,829)=5,66$; $p<0,001$). Ein Haupteffekt für Gruppe zeigte sich nicht. Es wurden signifikante Interaktionen zwischen Gruppe x Bedingung ($F(3,829)=7,04$, $p<0,001$) sowie ROI x Bedingung ($F(1,829)=2,96$, $p=0,02$) gefunden. Tabelle (3) fasst die sich aus den Interaktionen ergebenden Einzelanalysen für die vier experimentellen Gruppen zusammen. Dabei wird deutlich, dass nur die Gruppe 0-3 einen stabilen LAN Effekt evoziert, mit der stärksten Ausprägung im links-anterioren Bereich (siehe Abb. 3).

Gruppe	Bedingung		ROI		ROI x Bedingung	
	F	p	F	p	F	p
0 - 3	(1,223) = 34,10	< 0,0001 ***	(4,223) = 3,5	0,008 **	(4,223) = 4,34	0,002**
4 - 6	(1,194) = 2,71	n.s. (0,1)	(4,194) = 1,07	n.s. (0,3)	(4,194) = 0,32	n.s. (0,8)
7 - 10	(1,136) = 0,2	n.s. (0,65)	(4,136) = 1,89	n.s. (0,11)	(4,136) = 1,59	n.s. (0,17)
11 - 13	(1,252) = 5,67	0,02 *	(4,252) = 3,17	0,01 *	(4,252) = 0,9	n.s. (0,43)

Tabelle 3: statistische Analyse der jeweiligen Gruppen im Zeitfenster 300-500 ms

P600: Die statistische Analyse im Zeitfenster 600-800 ms zeigt signifikante Haupteffekte für Bedingung ($F(1,945)=244,61$, $p<0,0001$) und ROI ($F(4,945=10,69$, $p<0,0001$). Ein Haupteffekt für Gruppe wurde nicht gefunden. Die Berechnung ergab weiterhin signifikante Interaktionen zwischen ROI x Bedingung ($F(4,945)=6,13$; $p<0,0001$) sowie Gruppe x Bedingung ($F(3,945)=10,67$, $p<0,0001$). Tabelle (4) zeigt die sich daraus ergebenden Analysen einzeln für die jeweiligen experimentellen Gruppen.

Gruppe	Bedingung		ROI		ROI x Bedingung	
	F	p	F	p	F	p
0 - 3	(1,223) = 95,36	<0,0001 ***	(4,223)=5,25	0,005 **	(4,223) = 3,14	0,01 *
4 - 6	(1,252) = 60,56	<0,0001 ***	(4,252)=2,43	0,05 *	(4,252) = 2,11	n.s.(0,08)
7 - 10	(1,194) = 62,79	<0,0001 ***	(4,194)=1,83	n.s.(0,12)	(4,194) = 0,8	n.s. (0,51)
11 - 13	(1,252) = 27,77	<0,0001 ***	(4,252)= 4,1	0,003 **	(4,252) = 1,18	n.s. (0,31)

Tabelle 4: statistische Analyse der jeweiligen Gruppen im Zeitfenster 600-800 ms

Jede Gruppe zeigt einen signifikanten Effekt für Bedingung. Außer in der Gruppe 7-10 zeigt sich auch in jeder Gruppe ein signifikanter Effekt für ROI, indem die stärkste Ausprägung der P600 parietal zu finden ist (siehe Abb. 3). Eine Interaktion zwischen Bedingung x ROI wurde nur in der Gruppe 0-3 gefunden, ein Trend zu einer Interaktion zeigt sich auch in der Gruppe 4-6.

5. Diskussion

In dieser Studie wurde getestet, ob sich Unterschiede bei der Verarbeitung von Fehlern in der Kasusmarkierung am direkten Akkusativobjekt in der L2 aufgrund unterschiedlichen Erwerbsalters zeigen. Im Vordergrund stand dabei die Zeitspanne des kindlichen L2-Erwerbs, der hier auf einen Bereich zwischen vier und sechs Jahren festgelegt wurde. Die Daten zeigen unterschiedliche Ergebnisse für die behavioralen Daten und die EKP-Daten. In Bezug auf die behavioralen Daten lassen sich weder für Akkuratheit noch für die Reaktionszeiten quantitative Unterschiede zwischen simultanem und kindlichen L2-Erwerb finden. Beide Gruppen unterschieden sich weder in der Bewertung der Akkuratheiten untereinander, noch zwischen den Bedingungen. Auch in den Reaktionszeiten konnten weder Unterschiede zwischen den Gruppen, noch zwischen dem Verhalten innerhalb dieser beiden Gruppen gefunden werden. Es zeigt sich jedoch ein Alterseffekt im traditionellen Sinne (siehe 1), da sich die Ergebnisse der Gruppen 7-10 und 11-13 von denen der Gruppen 0-3 und 4-6 abgrenzen. Die Ergebnisse deuten darauf hin, dass sich ein später L2-Erwerb (> sechs Jahre) sowohl auf die Akkuratheit als auch auf die Reaktionszeit negativ bei der Beurteilung fehlerhafter Kasusmarkierung auswirkt. Die Akkuratheitsraten sind in den Gruppen mit spätem Erwerbsalter sowohl zwischen den Gruppen, als auch innerhalb der jeweiligen Gruppen zwischen den Bedingungen deutlich geringer. Lerner mit spätem Erwerbsalter (> sechs Jahre) bewerten fehlerhafte Strukturen weniger akkurat. Sie scheinen mehr Schwierigkeiten damit zu haben Inkorrektheiten zu erkennen und machen dementsprechend mehr Fehler in der Bewertung zwischen korrekten und inkorrekten Strukturen. Weiterhin waren die Gruppen mit spätem Erwerbsalter auch langsamer in den Reaktionszeiten. Dies geht einher mit der Annahme eines erhöhten Kapazitätenanspruches an die Sprachverarbeitung. Lerner mit spätem Erwerbsalter benötigen mehr Zeit um zu bewerten, ob eine bestimmte Struktur in der L2 korrekt ist, oder nicht.

Der Alterseffekt, der sich bezüglich der behavioralen Daten zeigt, kann als relativ robust angesehen werden, da u.a. der sprachliche Kenntnisstand als potenzielle Korrelationsquelle weitgehend kontrolliert wurde (siehe 3.2). Dieses Ergebnis bestätigt die zahlreichen Resultate bezüglich der Annahme eines Alterseffekts im L2-Erwerb, nämlich dass der Zeitpunkt eines Erwerbsalters um sechs Jahre für den L2-Erwerb einen Kritischen darstellt (u.a. Hyltenstam/Abrahamson: 2003).

Dagegen treten bezüglich der EKP-Daten unterschiedliche Ergebnisse hervor. Gemäß der untersuchten Komponenten (LAN und P600) zeigt sich, dass die Sprachverarbeitung im Sinne der automatischen Verarbeitung (Detektion) eines morphosyntaktischen Fehlers (LAN) und dessen Reanalyse / Integration (P600) unterschiedlich verläuft für Lerner mit simultanem als mit kindlichem und späten L2-Erwerb. Nur die Gruppe mit simultanem L2-Erwerb (Gruppe 0-3) evoziert ein stabiles LAN-P600 Verarbeitungsmuster. Hier wird die LAN-Komponente evoziert – also eine Reaktion darauf, dass eine fehlerhafte Markierung des Akkusativs vorliegt – (ähnlich) wie man sie auch bei monolingualen Sprechern erwarten würde. Die Lernergruppen 4-6 und 7-10 evozieren die LAN-Komponente nicht. Dies lässt darauf schließen, dass die Ergebnisse der Gruppe mit kindlichem L2-Erwerb ähnlich denen der Gruppe 7-10 sind und sich von denen der Gruppe mit simultanem L2-Erwerb unterscheiden. Die Ergebnisse bezüglich der LAN für die Gruppe 11-13 (siehe Tabelle 2) sind schwer zu deuten, da die visuelle Inspektion (siehe Abb. 3) keinen Hinweis auf eine Negativierung in diesem Zeitfenster (400-500 ms) gibt, sondern eher eine Positivierung aufweist. Eine Interpretation diesbezüglich ist schwierig und wäre höchst spekulativ. Auch die P600-Komponente findet sich stabil nur in der Gruppe 0-3. Die Gruppen mit kindlichem und spätem L2-Erwerb zeigen zwar robuste Unterschiede für Bedingung und Distribution, jedoch kann man hier aufgrund der fehlenden Interaktion nicht von einem klaren P600-Korrelat sprechen. Es scheint, dass auch in Bezug auf die späten kontrollierten Verarbeitungsprozesse das Erwerbsalter bereits früh einen Einfluss hat. Auch hier bezieht sich die Erklärung für die Verarbeitungsunterschiede auf die erhöhten Ansprüche an die Verarbeitungskapazität. Diese werden bereits sichtbar, sobald das L2-Erwerbsalter im Bereich des kindlichen L2-Erwerbs liegt, wenn auch nicht so klar, wie für frühe automatisierte Verarbeitungsprozesse (z.B. LAN).

Zusammenfassend ist festzuhalten, dass die hier berichteten EKP-Daten die Ergebnisse bisheriger Studien größtenteils unterstreichen. Es konnte gezeigt werden, dass vor allem frühe syntaktische Korrelate wie die LAN anfällig für einen Alterseffekt im L2-Erwerb sind (Hahne: 2001; Müller/Rüschemeyer/Friederici: 2006). Dies geht einher mit den bisherigen Annahmen, dass mit steigendem L2-Erwerbsalter höhere Ansprüche an die Kapazitäten des Sprachprozessors, z.B. bei der Detektion fehlerhafter Kasusmarkierung, gestellt werden. Diese höhere Beanspruchung wirkt sich bereits dann aus, wenn die L2 nicht simultan erworben wurde. Es ist dementsprechend anzunehmen, dass bereits ab dem Zeitpunkt des kindlichem L2-Erwerb L2-Lerner einen erhöhten Kapazitätenanspruch erheben. Dies geht einher mit der Annahme, dass Erwerbsprozesse der Kasusmarkierung bereits früh im (Mutter-)Spracherwerb abgeschlossen sind und das volle Potenzial der diesbezüglichen höchst automatisierten Verarbeitungsprozesse wohl nur bei Muttersprachlern zu finden ist. Defizite in der Verarbeitung von fehlerhafter Kasuszuweisung zeigen sich demnach bereits bei Lernern mit einem Erwerbsalter zwischen vier und sechs Jahren aufgrund verminderter Automatisierung derer Verarbeitungsprozesse (vgl. Hahne: 2001). Auch in Bezug auf die späten Verarbeitungskomponenten, wie die P600, zeigen die Ergebnisse dieser Studie einen Zusammenhang mit dem Erwerbsalter, wenn auch nicht so klar, wie die LAN betreffend. In jeder Gruppe trat ein Bedingungseffekt hervor, jedoch wurde eine klare P600-Komponente nur von der Gruppe 0-3 evoziert. Dies bestätigt die bisherigen Ergebnisse nur teilweise, welche späte Verarbeitungsprozesse wie die P600 weniger stark mit einem Alterseffekt in Verbindung bringen, als die frühen Verarbeitungsprozesse, da es sich um Prozesse handelt, die stärker kontrolliert ablaufen und weniger auf Automatisierung angewiesen sind. Die Annahme, die dem zu Grunde liegt ist, dass Lerner mit einem hohen Kenntnisstand in ihrer L2 Prozesse für kontrollierte Verarbeitung erworben haben, bzw. diese auch für die L2 abrufen können.

Diese Studie deutet an, dass es Unterschiede in der L2-Sprachverarbeitung bei fehlerhafter Kasusmarkierung gibt und dass diese Unterschiede in Zusammenhang mit dem L2-Erwerbsalter stehen. Es konnte weiterhin gezeigt werden, dass die Gruppe mit kindlichem L2-Erwerb Unterschiede zur Gruppe mit simultanem L2-Erwerb und mehr Gemeinsamkeiten mit der Gruppe 7-10 (mit spätem Erwerbsalter) aufweist, dieses Ergebnis wurde jedoch nur in den EKP-Daten gefunden. Die Ergebnisse der behavioralen Daten weisen eher darauf hin, dass der Einfluss des Erwerbsalters erst spät mit ab ca. sechs Jahren hervor tritt. Anders als bei Weber-Fox und Neville (1996) deuten die Ergebnisse hier in Bezug auf den kindlichen L2-Erwerb eher Gemeinsamkeiten zum simultanen und Unterschiede zum späten L2-Erwerb an. Ein Grund dafür könnte sein, dass die hier untersuchte syntaktische Verletzung (fehlerhafte Kasusmarkierung) salient war, während die bei Weber-Fox und Neville untersuchten Strukturen eher Verletzungen (u.a. Subjazenz) hervorriefen, die weniger salient waren. Die EKP-Methode scheint für Untersuchungen dieser Art gegenüber Reaktionszeit- und Akkuratheitsmessungen den Vorteil zu haben, feinere Verarbeitungsunterschiede evozieren zu können. Es braucht zukünftig mehr Studien dieser Art, um diesbezüglich robuste Aussagen über den Bereich des kindlichen L2-Erwerbs zu machen.

Literatur

Baayen, R. H. (2008), *Analyzing Linguistic Data*, Cambridge: Cambridge University Press.
Babyonyshev, M. (1993), „The acquisition of Russian case", *MIT Working Papers in Linguistics*, 19, Cambridge, MA: MIT Press, 1-44.
Bates, D. / M. Maechler / B. Bolker (2011), *Linear mixed-effects models using S4 classes*, R package version 0.999375-40: http://CRAN.R-project.org/package=lme4. [Zugriff am 14 Juni 2011]
Eisenbeiss, S. / S. Bartke / H. Clahsen (2006), „Structural and Lexical Case in Child German: Evidence from language-impaired and typically-developing children", *Language Acquisition*, 13, 3-32.
Eisenbeiss, S. / B. Narashiman / M. Voeikova (2009), „The acquisition of case", in: Malchukov A. / A. Spencer (Hgg.), *The Oxford Handbook of Case*, Oxford: Oxford University Press, 369-382.
Friederici, A. D. (1995), „The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data", *Brain and Language*, 50, 259-281.
Friederici, A. D. (2002), „Towards a neural basis of auditory sentence processing", *Trends in Cognitive Sciences*, 6, 78-84.
Friederici, A. D. / A. Hahne / D. Saddy (2002), „Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair", *Journal of Psycholinguistic Research*, 31, 45-63.
Hahne, A. (2001), „What's different in second-language processing? Evidence from event-related brain potentials", *Journal of Psycholinguistic Research*, 30, 251-266.
Hyltenstam, K. / N. Abrahamsson (2003), „Maturational constraints in SLA", in: Doughty, C. J. / M. H. Long (Hgg.), *The Handbook of Second Language Acquisition*, Oxford: Blackwell, 539-588.
Kaan, E. / A. Harris / E. Gibson / P. J. Holcomb (2000), „The P600 as index of syntactic integration difficulty", *Language and Cognitive Processes*, 12, 159-201.
Kliegl, R. / P. Wei / M. Dambacher / M. Yan / X. Zhou (2011), „Experimental effects and individual differences in linear mixed models: estimating the relationship between spatial, object and attraction effects in visual attention", *Frontiers in Quantitative Psychology and Measurement*, 1, 1-12.
Lenneberg, E. H. (1967), *Biological foundations of language*, New York: Wiley & Sons.
Long, M. H. (1990), „Maturational constraints on second language acquisition", *Studies in Second Language Acquisition*, 12, 251-285.
Meisel, J. M. (2008), „Child second language acquisition or successive first language acquisition", in: Haznedar, B. / E. Gavruseva (Hgg.), *Current trends in child second language acquisition: A generative perspective*, Amsterdam: John Benjamins, 55-82.
Meisel, J. M. (2010), „Age of onset in successive acquisition of bilingualism", in: Kail, M. / M. Hickman (Hgg.), *Language Acquisition across Linguistic and Cognitive Systems*, Amsterdam: John Benjamins, 225-247.
Müller, J. L. / S. A. Rüschemeyer / A. D. Friederici (2006), „Aktivitätsmuster im Gehirn: Unterschiede und Gemeinsamkeiten beim Verstehen von Erst- und Zweitsprache", *Neuroforum*, 2, 176-184.
Müller, J. L. / M. Hirotani / A. D. Friederici (2007), „ERP evidence for different strategies in the processing of case markers in native speakers and non-native learners", *BMC Neuroscience (Online Journal)*, 8.
Oldfield, R. C. (1971), „The assessment and analysis of handedness: The Edinburgh inventory", *Neuropsychologia*, 9, 97-113.
Osterhout, L. / P. J. Holcomb (1992), „Event-related brain potentials elicited by syntactic anomaly", *Journal of Memory and Language*, 31, 785-806.

Osterhout, L. / J. McLaughlin / I. Pitkanen / C. Frenck-Mestre / N. Molinaro (2006), „Novice learners, longitudinal designs, and event-related potentials: A paradigm for exploring the neurocognition of second-language processing", *Language Learning*, 56, 199-230.
Radford, A. (1990), *Syntactic theory and the acquisition of English syntax: the nature of early child grammars of English*, Oxford: Blackwell.
Rothweiler, M. (2007), „Bilingualer Spracherwerb und Zweitspracherwerb", in: Steinbach, M. / R. Albert / H. Girnth / A. Hohenberger / B. Kümmerling-Meibauer / J. Meibauer / M. Rothweiler / M. Schwarz-Friesel (Hgg.), *Schnittstellen der germanistischen Linguistik*, Stuttgart / Weimar: J. B. Metzler, 103-135.
Sharbrough, F. W. / G. E. Chatrian / R. P. Lesser / H. O. Lüders / M. R. Nuwer / T. W. Picton (1991), „American Electroencephalographic Society guidelines for standard electrode position nomenclature", *Journal of Clinical Neurophysiology*, 8, 200-202.
Singleton, D. (2005), „The critical period hypothesis: A coat of many colours", *IRAL*, 46, 269-285.
Weber-Fox, C. M. / H. J. Neville (1996), „Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers", *Journal of Cognitive Neuroscience*, 8, 231-256.
Wegener, H. (1995). *Die Nominalflexion des Deutschen – verstanden als Lerngegenstand*, Tübingen: Niemeyer.

7 Appendix: Mail Questionnaire on Handedness and Language Background Based on the Edinburgh Handedness Design (Oldfield, 1971): Native Speakers

Name: **Datum:** **Alter:**

Geschlecht: **Studienfach:** **Semester:**

Mit welcher Hand führen Sie die aufgeführten Tätigkeiten bevorzugt aus?

x = bevorzugt xx = immer unsicher = x in beiden Spalten

	linke Hand	rechte Hand
Schreiben		
Zeichnen		
Werfen		
Schneiden		
Zähneputzen		
Messer (ohne Gabel)		
Löffel		
Besen (obere Hand)		
Schlaghand (beim Spiel)		
Behälter öffnen		

Gibt es in Ihrer Familie Linkshänder/innen? Bitte kreuzen Sie an (x):

	ja	nein	weiß ich nicht
Mutter			
Vater			
Großmutter (mütterlicherseits)			

Großmutter (väterlicherseits)			
Großvater (mütterlicherseits)			
Großvater (väterlicherseits)			
Geschwister 1			
Geschwister 2			
Geschwister 3			

Wie schätzen Sie Ihre Deutschkenntnisse auf einer Skala von 0 – 5 ein?
0 = keine Kenntnisse; 1 = Grundkenntnisse; 2 = mittelmäßig; 3 = flüssig; 4 = sehr flüssig; 5 = muttersprachlich

Hörverstehen	
Leseverstehen	
Sprechen	
Schreiben	

Für die Organisation von Sprache im Gehirn spielt es möglicherweise eine Rolle, ob man weitere Sprachen gelernt hat bzw. mehrere Muttersprachen besitzt. Deshalb möchten wir Sie bitten, in die Tabelle unten einzutragen, welche Sprachen Sie noch gelernt haben und wie gut.

Beispiel:

Sprache	**Kenntnisstand**	**wann gelernt? (Alter)**	**wo/wie gelernt?**	**Verwendung**
Deutsch	Muttersprache			
Englisch		10 Jahre	Schule & Ausland (Schüleraustausch)	Schule & Kommuni-kation im Alltag
Sprache	**Kenntnisstand**	**wann gelernt? mit welchem Alter**	**wo/wie gelernt?**	**Verwendung**
Deutsch	Muttersprache			
Englisch				
Französisch				
Russisch				
Spanisch				
weitere:				

Möglicherweise kann auch der sprachliche Hintergrund die sprachliche Verarbeitung beeinflussen, daher noch ein paar letzte Fragen:

Geburtsort/ Region:	
In welcher Region sind Sie aufgewachsen?	
Welche Muttersprache haben Ihre Eltern?	
Mutter	
Vater	
In welcher Sprache kommunizieren Sie mit ihrer/m/n:	
Mutter	
Vater	
Geschwistern	
Großeltern (mütterlicherseits)	
Großeltern (väterlicherseits)	
Freunden (Freizeit)	
Wie lange leben Sie schon in Berlin/ Brandenburg?	
Wie oft benutzen Sie Ihre Muttersprache in den folgenden Situationen:	**0 = niemals; 1 = nur wenn nötig; 2 = selten; 3 = manchmal; 4 = meistens; 5 = immer**
in der Uni/ auf der Arbeit	
im Alltag (einkaufen, etc.)	
zu Hause	
mit Freunden (hauptsächlich)	
Haben Sie Kinder? Wenn ja, in welcher Sprache kommunizieren Sie mit ihnen?	
Welchen Schulabschluss haben Sie?	
Haben Sie nach dem Schulabschluss eine Ausbildung absolviert? Wenn ja, welche?	

Vielen Dank!

Hiermit versichere ich Ihnen, dass die hier von Ihnen gemachten Angaben anonym behandelt werden und rein wissenschaftlichen Zwecken dienen. Ihre Daten werden nicht weitergeleitet und nur zum Forschungszwecke dieser Studie verwendet!

8 Appendix: Mail Questionnaire on Handedness and Language Background Based on the Edinburgh Handedness Design (Oldfield, 1971): L2 learners

Name: **Datum:** **Alter:**

Geschlecht: **Studienfach:** **Semester:**

Mit welcher Hand führen Sie die aufgeführten Tätigkeiten bevorzugt aus?
x = bevorzugt xx = immer unsicher = x in beiden Spalten

	linke Hand	rechte Hand
Schreiben		
Zeichnen		
Werfen		
Schneiden		
Zähneputzen		
Messer (ohne Gabel)		
Löffel		
Besen (obere Hand)		
Schlaghand (beim Spiel)		
Behälter öffnen		

Gibt es in Ihrer Familie Linkshänder/innen? Bitte Kreuzchen (x) machen:

	ja	nein	weiß ich nicht
Mutter			
Vater			
Großmutter (mütterlicherseits)			
Großmutter (väterlicherseits)			

Großvater (mütterlicherseits)			
Großvater (väterlicherseits)			
Geschwister 1			
Geschwister 2			
Geschwister 3			

Wie schätzen Sie Ihre Deutschkenntnisse auf einer Skala von 0 – 5 ein?
0 = keine Kenntnisse; 1 = Grundkenntnisse; 2 = mittelmäßig; 3 = flüssig; 4 = sehr flüssig; 5 = muttersprachlich

Hörverstehen	
Leseverstehen	
Sprechen	
Schreiben	

Für die Organisation von Sprache im Gehirn spielt es möglicherweise eine Rolle, ob man weitere Sprachen gelernt hat bzw. mehrere Muttersprachen besitzt. Deshalb möchten wir Sie bitten, in die Tabelle unten einzutragen, welche Sprachen Sie noch gelernt haben und wie gut.

Beispiel:

Sprache	**Kenntnisstand**	**wann gelernt? (Alter)**	**wo/wie gelernt?**	**Verwendung**
Polnisch	Muttersprache			
Deutsch		10 Jahre	Schule	nur Schule
Englisch		16 Jahre	Schule & Ausland (Schüleraustausch)	Schule & Kommunikation im Alltag

Sprache	**Kenntnisstand**	**wann gelernt? Alter**	**wo/wie gelernt?**	**Verwendung**
Polnisch				
Deutsch				
Englisch				
Russisch				
Französisch				
weitere:				

Möglicherweise kann auch der sprachliche Hintergrund die sprachliche Verarbeitung beeinflussen, daher noch ein paar letzte Fragen:

Geburtsort/ Region:	
In welcher Region sind Sie aufgewachsen?	
Welche Muttersprache haben Ihre Eltern?	
Mutter	
Vater	
In welcher Sprache kommunizieren Sie mit ihrer/m/n:	
Mutter	
Vater	
Geschwistern	
Großeltern (mütterlicherseits)	
Großeltern (väterlicherseits)	
Freunden (Freizeit)	
Wie lange leben Sie schon in Berlin/ Brandenburg? (Angabe in Jahren)	
Wie oft benutzen Sie die deutsche Sprache: (Skala 0-5)	**0 = niemals; 1 = nur wenn nötig; 2 = selten; 3 = manchmal; 4 = meistens; 5 = immer**
in der Uni/ auf der Arbeit	
im Alltag (einkaufen, etc.)	
zu Hause (Familie)	
mit Freunden	
Haben Sie Kinder? Wenn ja, in welcher Sprache kommunizieren Sie mit ihnen?	
Welchen Schulabschluss haben Sie?	
Haben Sie nach dem Schulabschluss eine Ausbildung absolviert? Wenn ja, welche?	

Vielen Dank!

Hiermit versichere ich Ihnen, dass die hier von Ihnen gemachten Angaben anonym behandelt werden und rein wissenschaftlichen Zwecken dienen. Ihre Daten werden nicht weitergeleitet und nur zum Forschungszwecke dieser Studie verwendet!

9 Appendix: C-Test

Humboldt Universität zu Berlin

VP Nr.:

Gruppe: Datum:

In den folgenden Texten fehlt in jedem zweiten Wort der letzte Teil des Wortes. Bitte ergänzen Sie die fehlenden Teile (in Blockschrift)! Sie haben dafür 30 Minuten Zeit.
Umlaute (ä. ö. ü) und ß gelten jeweils als ein Buchstabe.

Familienkonferenz macht Kinder selbstbewusst

Familienkonferenzen können Kindern zu mehr Selbstbewusstsein verhelfen. Beim gleichber_ _ _ _ _ _ _ _ _ Austausch m_ _ anderen Familien-m_ _ _ _ _ _ _ _ _ _ üben Kin_ _ _ , sich f_ _ ihre eig_ _ _ _ Probleme u_ _ Interessen st_ _ _ zu mac_ _ _ . Kleinere Kin_ _ _ sollten da_ _ _ genauso vi_ _ _ Mitspracherechte ha_ _ _ wie d_ _ älteren Gesch_ _ _ _ _ _ oder Elt_ _ _ . Jeder h _ _ die glei_ _ _ _ Rechte – d_ _ 4-jährige Ki_ _ und der 40-jährige Familienvater stehen auf der gleichen Stufe. Wie oft eine solche Konferenz abgehalten wird, hängt laut Professor Struck, Erziehungswissenschaftler von der Universität Hamburg, von den Bedürfnissen der jeweiligen Familie ab.

O.K.! oder Okay! oder …?

Wie oft sagt man am Tag: Alles „Okay". Unschwer zu hören ist, dass das Wort aus dem Englischen, genauer gesagt aus dem amerikanischen Englisch in die deutsche Sprache übernommen wurde. Wie ab_ _ sind d _ _ Amerikaner zu d_ _ _ _ _ Wendung geko_ _ _ _? Sprachforscher ha_ _ _ fast 30 Lösu_ _ _ _ gefunden. So ge_ _ die Bezei_ _ _ _ _ _ zum Beis_ _ _ _ auf schl _ _ _ _ _ Englischkenntnisse deut_ _ _ _ _ Einwanderer zur _ _ _ . Ei _ _ andere Erkl _ _ _ _ _ beschreibt, da_ _ afrikanische Skl_ _ _ _ das Wo_ _ nach Ame_ _ _ _ mitbrachten. „Oke" bedeutet in der westafrikanischen Mandigo-Sprache „in Ordnung".

domke 1

Humboldt Universität zu Berlin

Schulanfänger: Die Eltern sind gefragt

Zahlreiche Kinder in der Bundesrepublik Deutschland können die Schule nicht beginnen, weil sie Probleme haben. Es fehlt ihnen an Reife, sie kön_ _ _ sich ni _ _ _ konzentrieren u _ _ immer öf_ _ _ sind ih_ _ Kenntnisse d _ _ Muttersprache schl_ _ _ _ . Die Elt_ _ _ tragen ei_ _ sehr gr_ _ _ Verantwortung da_ _ _ , ob i_ _ Kind re_ _ für d_ _ Unterricht i_ _ . So ei _ _ Vorbereitung gel_ _ _ _ natürlich ni_ _ _ , wenn z_ _ Beispiel der Fern_ _ _ _ _ die Erziehung der Kinder übernimmt. Mutter und Väter müssen bei ihren Kindern Ausdauer, Motivation und Konzentration fördern. Die Realität sieht leider oft anders aus. Nach einer Umfrage sprechen Kinder und Eltern immer weniger miteinander.

Scheidung verdoppelt Armutsrisiko der Frauen

Nach einer Trennung oder Scheidung versorgt meistens die Mutter die Kinder. 25 Pro_ _ _ _ der Fra _ _ _ in Deuts _ _ _ _ _ _ erhalten kei_ _ _ Kinder-unterhalt, w_ _ eine v_ _ der Unive_ _ _ _ _ _ Bielefeld veröffe_ _ _ _ _ _ _ _ Studie er_ _ _ . Diese Stu_ _ _ untersuchte d_ _ wirtschaftlichen Fol_ _ _ von Sche_ _ _ _ _ und Tren_ _ _ _ . Scheidung u_ _ Trennung verdo_ _ _ _ _ das Armutsr_ _ _ _ _ der Fra_ _ _ . Trotzdem wi _ _ die Sche_ _ _ _ _ in zwei Dritteln der Fälle von den Frauen beantragt.

Humboldt Universität zu Berlin

Picasso-Museum als Malagas Magnet Es ist die neue Attraktion von Malaga. Im Oktober 2003 öffnete in der Geburtsstadt Pablo Picassos ein neues Museum, das me_ _ als 200 Gem_ _ _ _ , Zeichnungen u_ _ Skulpturen d _ _ einflussreichsten Küns_ _ _ _ _ des 20. Jahrhu _ _ _ _ _ _ zeigt. D_ _ Werke sta_ _ _ _ aus d_ _ Besitz sei_ _ _ Erben. D_ _ Museum befi_ _ _ _ sich in ei_ _ _ Palast d_ _ Grafen v _ _ Buenavista a_ _ dem 16. Jahrh_ _ _ _ _ _ . Es erwa_ _ _ _ - Schulkassen ni_ _ _ mitgerechnet - jähr_ _ _ _ 1,4 Millionen Besucher, davon gut 300 000, die allein in die südspanische Stadt reisen. Planungen gehen davon aus, dass die Kunsttouristen pro Jahr mehr als 30 Millionen Euro nach Malaga bringen.	

10 Appendix: Approved Consent Form for Participation in an EEG Study

EINWILLIGUNGSERKLÄRUNG
Für Teilnehmer/innen an der Studie

Studientitel: ***EKP – L2POL – GSV***

Hiermit erkläre ich ____

(Vorname, Name, Adresse des/der Teilnehmers/in)

geb. am_________________, dass ich durch____*Juliane Domke*________ mündlich und schriftlich über das Wesen, die Bedeutung und Risiken der wissenschaftlichen Untersuchung im Rahmen der o.g. Studie informiert wurde und ausreichend Gelegenheit hatte, meine Fragen hierzu in einem Gespräch zu klären.

Ich habe insbesondere die mir vorgelegte TeilnehmerInneninformation verstanden und eine Ausfertigung derselben und dieser Einwilligungserklärung erhalten.

Mir ist bekannt, dass ich meine Einwilligung jederzeit ohne Angabe von Gründen und ohne nachteilige Folgen für mich zurückziehen und einer Weiterverarbeitung meiner Daten jederzeit widersprechen und ihre Löschung bzw. Vernichtung verlangen kann.

Ich bin bereit an der wissenschaftlichen Untersuchung im Rahmen der o.g. Studie teilzunehmen.

Ich erkläre mich einverstanden, dass die im Rahmen dieser Studie erhobenen Daten meiner Person verschlüsselt auf elektronischen Datenträgern aufgezeichnet und verarbeitet und die anonymisierten Studienergebnisse veröffentlicht werden.

Ich habe jederzeit das Recht, Fragen, welche die Studie betreffen, an die/den Verantwortliche/n zu stellen. Fragen diesbezüglich sind zu richten an: Juliane Domke, M.A., Tel.: 030-20939682.

Berlin, den _______________ ______________________________

(Unterschrift des/der Teilnehmers/in)

Appendix 10

Hiermit erkläre ich, den/die Teilnehmer/in über über das Wesen, die Bedeutung und Risiken der wissenschaftlichen Untersuchung im Rahmen der o.g. Studie aufgeklärt und ihm/ihr eine Ausfertigung der Information sowie dieser Einwilligungserklärung übergeben zu haben.

Berlin, den ______________ ______________________________

(Unterschrift des/der Versuchsleiters/in)

11 Appendix: Post-Questionnaire on Self-Perceived Difficulty to Accomplish the Experimental Task

VP-Nr.:

Datum:

1. Wie beurteilen Sie die Dauer des Experimentes?

sehr kurz mittel sehr lang

2. Wie empfanden Sie die Präsentation der Sätze?

sehr langsam mittel sehr schnell

3. Die Aufgabe war für Sie:

sehr leicht mittel sehr schwer

4. Wie sind Sie vorgegangen, um die Aufgabe zu lösen?

5. Was glauben Sie, wollen wir mit diesem Experiment untersuchen?